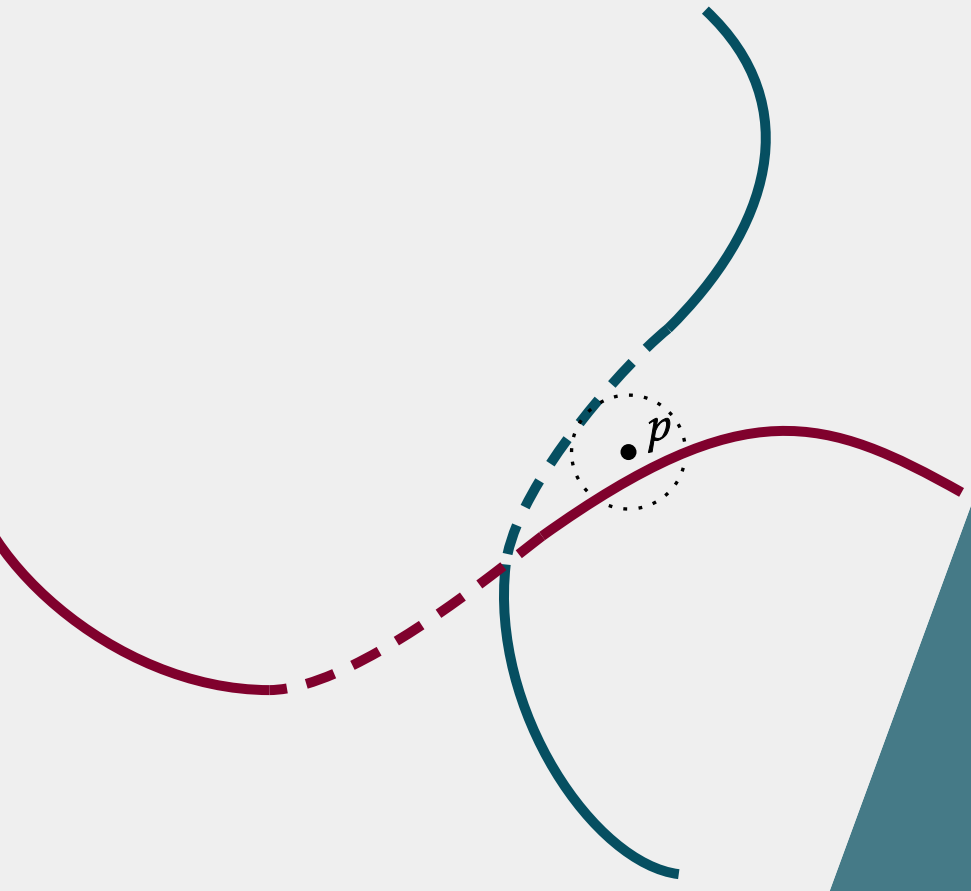
Algebraic and semi-algebraic phylogenetic reconstruction

Marina Garrote-López



UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

ALGEBRAIC AND SEMI-ALGEBRAIC PHYLOGENETIC RECONSTRUCTION

by

Marina GARROTE-LÓPEZ

PhD in Applied Mathematics

Marta CASANELLAS, advisor
Jesús FERNÁNDEZ-SÁNCHEZ, advisor

Barcelona, June 2021

Als meus pares

Abstract

Phylogenetics is the study of the evolutionary history and relationships among groups of biological entities (called *taxa*). The modeling of those evolutionary processes is done by phylogenetic trees whose nodes represent different taxa and whose branches correspond to the evolutionary processes between them. The leaves usually represent contemporary taxa and the root is their common ancestor. Nowadays, phylogenetic reconstruction aims to estimate the phylogenetic tree that best explains the evolutionary relationships of current taxa using solely information from their genome arranged in an alignment. We focus on the reconstruction of the topology of phylogenetic trees, which means reconstructing the shape of the tree considering labels at the leaves.

To this end, one usually assumes that DNA sequences evolve according to a Markov process ruled by a prescribed model of nucleotide substitutions. These substitution models specify transition matrices associated to the edges of the tree and a distribution of nucleotides at the root. Given a tree T and a model of nucleotide substitutions, one can compute the distribution of nucleotide patterns at the leaves of T in terms of the parameters of the model. This joint distribution is represented by a vector whose entries correspond to the joint probability of every possible nucleotide observation. Those entries can be expressed as polynomials on the model parameters and satisfy certain algebraic relationships. The study of these relationships and the geometry of the algebraic varieties defined by them (called *phylogenetic varieties*) have provided successful insight into the problem of phylogenetic reconstruction. However, from a biological perspective we are not interested in the whole variety, but only in the region of points that arise from stochastic parameters (the so-called *phylogenetic stochastic region*). The description of these regions leads to semi-algebraic constraints which play an important role since they characterize distributions with a biological and probabilistic meaning. One of the main motivations for this thesis follows from the following question. *Could the use of semi-algebraic tools improve the already existent algebraic tools for phylogenetic reconstruction?*

To answer this question, we compute the Euclidean distance of data points arising from an alignment of nucleotide sequences to the phylogenetic varieties and their stochastic regions. We deal with a number of scenarios of special interest in phylogenetics, such as trees with short external branches and/or subject to the long branch attraction phenomenon. In some cases, we compute these distances analytically and we can decide which tree has stochastic region closer to the data point. As a consequence, we can prove that, even if the data point was close to the phylogenetic variety of a given tree, it might be closer to the stochastic region of another tree. In particular, considering the stochastic phylogenetic region seems to be fundamental to cope with the phylogenetic reconstruction problem when dealing with the long branch attraction phenomenon.

However, incorporating semi-algebraic tools into phylogenetic reconstruction methods can be extremely difficult and the procedure to do it is not evident at all. In this thesis, we present two phylogenetic reconstruction methods that combine algebraic and semi-algebraic conditions for the general Markov model. The first method we present is `SAQ`, which stands for *Semi-Algebraic Quartet reconstruction method*. Next, we introduce a more versatile method, `ASAQ` (for *Algebraic and Semi-Algebraic Quartet reconstruction method*), which combines `SAQ` with the method `Erik+2` (based on certain algebraic constraints). Both are phylogenetic reconstruction methods for DNA alignments on four taxa which have been proven to be statistically consistent.

We test the suggested methods on simulated and real data to check their actual performance in several scenarios. Our simulation studies show that both methods `SAQ` and `ASAQ` are highly successful, even when applied to short alignments or data that violates their assumptions.

Resum

La filogenètica és l'estudi de la història evolutiva i les relacions entre grups d'entitats biològiques (anomenades *tàxons*). Aquests processos evolutius estan modelitzats per arbres filogenètics, els nodes dels quals representen diferents tàxons i les branques corresponen als processos evolutius entre ells. Les fulles normalment representen tàxons actuals i l'arrel és el seu avantpassat comú. Actualment, la reconstrucció filogenètica pretén estimar l'arbre filogenètic que millor explica les relacions evolutives de tàxons actuals utilitzant únicament informació del seu genoma organitzada en un alineament. En aquesta tesi ens centrem en la reconstrucció de la topologia dels arbres filogenètics, és a dir, reconstruir la forma de l'arbre tenint en compte els noms associats a les fulles.

Amb aquesta finalitat, assumim que les seqüències d'ADN evolucionen segons un procés de Markov d'acord amb un model de substitució de nucleòtids. Aquests models de substitució assignem matrius de transició a les arestes d'un arbre i una distribució de nucleòtids a l'arrel. Donat un arbre i un model de substitucions de nucleòtids, es pot calcular la distribució de les possibles observacions de nucleòtids a les seves fulles en termes dels paràmetres del model. Aquesta distribució conjunta s'acostuma a representar amb un vector, les entrades del qual es poden expressar com polinomis en els paràmetres del model i satisfan certes relacions algebraiques. L'estudi d'aquestes relacions i de la geometria de les varietats algebraiques que defineixen (anomenades *varietats filogenètiques*) han servit per entendre millor el problema de la reconstrucció filogenètica. No obstant això, des d'una perspectiva biològica no estem interessats en tota la varietat, sinó només en la regió de punts que resulten de paràmetres estocàstics (l'anomenada *regió estocàstica*). La descripció d'aquestes regions condueix a restriccions semialgebraiques que tenen un paper important, ja que caracteritzen les distribucions amb significat biològic i probabilístic. Una de les principals motivacions d'aquesta tesi és la següent pregunta: *Podria l'ús d'eines semialgebraiques millorar les eines algebraiques ja existents per a la reconstrucció filogenètica?*

Per respondre a aquesta pregunta, calculem la distància euclidiana entre punts de dades obtinguts a partir d'un alineament i varietats filogenètiques i les seves regions estocàstiques. Estudiem en una sèrie d'escenaris d'especial interès en la filogenètica, com ara arbres amb branques externes curtes i/o subjectes al fenomen d'atracció de branques llargues. En alguns casos, podem calcular aquestes distàncies de forma analítica i això ens permet demostrar que, fins i tot si el punt de dades fos proper a la varietat filogenètica d'un arbre donat, podria estar més a prop de la regió estocàstica d'un altre arbre. En particular, considerar la regió estocàstica sembla ser fonamental per fer front al problema de la reconstrucció filogenètica quan tractem amb del fenomen d'atracció de branques llargues.

No obstant això, la incorporació d'eines semialgebraiques en els mètodes de reconstrucció filogenètica pot ser extremadament difícil i el procediment per fer-ho no és gens evident. En aquesta tesi, presentem dos mètodes de reconstrucció filogenètica que combinen condicions algebraiques i semialgebraiques per al model general de Markov. El primer mètode que presentem és el `SAQ`, que rep el nom de *Semi-Algebraic Quartet reconstruction method*. A continuació, introduïm un mètode més versàtil, l'`ASAQ` (*Algebraic and Semi-Algebraic Quartet reconstruction method*), que combina el `SAQ` amb el mètode `Erik+2` (basat en certes restriccions algebraiques). Tots dos són mètodes de reconstrucció filogenètica per a alineaments d'ADN per quatre tàxons i hem demostrat que els dos són estadísticament consistents.

Finalment, testem els mètodes proposats amb dades simulades i dades reals per comprovar el seu rendiment en diversos escenaris. Les nostres simulacions mostren que ambdós mètodes SAQ i ASAQ obtenen molt bons resultats, fins i tot quan s'utilitzen amb alineaments curts o amb dades inconsistents amb les hipòtesis teòriques en que es basen.

Acknowledgements

En primer lloc vull donar les gràcies als meus directors, la Marta i el Jesús, per la seva enorme dedicació, suport i paciència durant aquests últims anys, fins i tot abans de començar aquesta tesi. Gràcies per encoratjar-me a seguir endavant i per deixar-me la llibertat d'aprendre i d'explorar nous camins. Per totes les discussions (a vegades repetides), les explicacions i per resoldre'm qualsevol dubte que hagi pogut tenir. I moltíssimes gràcies per llegir i rellegir aquesta tesi. Una tesi comença molt abans de quan t'imagines. En el meu cas va començar ara fa 10 anys en una classe de la FME; gràcies Marta per prendre't la llibertat de dedicar una sessió d'àlgebra multilineal a explicar-nos què era la filogenètica i quina relació hi tenia amb l'àlgebra, en aquell moment vaig veure clar cap a on podia anar el meu futur. He sigut molt afortunada de tenir-vos com a directors.

Una de les coses més inesperades que m'ha aportat fer aquesta tesi és tota la gent que he conegut i amb qui he pogut compartir algun moment d'aquests últims anys. A tots ells els vull dedicar un sincer agraïment.

Al Jordi per tots els viatges que hem fet junts anant a seminaris i congressos arreu, pels bons moments i per totes les anècdotes que ens emportem. Gràcies Jordi per escoltar-me practicar les meves xerrades mil vegades, ha sigut un plaer tenir-te de "germà" de doctorat. Al Marc Sabaté per tenir els seus codis disponibles, ben documentats i per deixar-me utilitzar-los. Al Raül, perquè, tot i que fos breu, em va agradar tenir-lo a prop en aquesta recerca i per respondre (tot i no sempre ser recíproc) i escoltar-me sempre que ho he necessitat. I al Guillem per tot el suport tècnic i per les seves útils i eficients respostes als meus problemes.

I would also like to thank Elizabeth Allman and John Rhodes for their interest in this research, for inviting me to visit them in Alaska and for giving me the opportunity to work with them. Although this moment has not yet arrived, I hope it will come soon. Thanks also to Piotr Zwernik for listening to us and always coming up with interesting ideas.

Grazie mille Pietro e Alessandro! Me hace muy feliz que vuestra carrera os trajera hasta Barcelona. Gracias por todas las horas de trabajo compartidas, pero sobre todo gracias por las pausas y las largas conversaciones sobre comida, política, idiomas y muchos otros temas. Y Ale, gracias por darme siempre buenos consejos, guiarme, animarme a seguir adelante y alegrarte de mis progresos. Compartir con vosotros el despacho ha sido una de las mejores cosas de estos años. Also thank you Lin for your happiness, your generosity, for your enthusiasm to explain many things to us and for wanting to listen even more.

Quiero dedicar un especial agradecimiento a una perfecta "intersección completa". Beatriz, muchas gracias, moltes gràcies Roser. Me encanta veros cada semana ya sea para discutir matemáticas o para hablar del presente, del futuro o de las alegrías y frustraciones de esta profesión. Me alegro de haberos tenido cerca este último año, he aprendido muchísimo de vosotras y escucharos siempre me inspira y reconforta a partes iguales. Para mí, sois un ejemplo a seguir.

I si les persones que he conegut en aquest camí han sigut importants, encara ho han estat més els de sempre, que d'una manera o altra m'han acompanyat aquests últims anys. A tots ells, gràcies.

Al Xavi i al Miki pel millor coworking! Pels esmorzars contundents, els dinars sempre compartits i el polvorons de postres que mai han faltat. Gràcies pels dies de màxima concentració i productivitat i pels de molta dispersió, riures i grans converses. Gràcies Xavi per tot el teu suport, més enllà del coworking i de la tesi. Per preocupar-te per mi, ajudar-me quan ho necessito i fer-me entendre que sempre puc comptar amb tu, potser una petita part d'aquesta tesi també és teva. I a la Paula per enviar-me sempre l'emoticono adequat, per pressionar-me quan toca i donar-me espai quan ho necessito, per llegir tot el que li envio i per tot el carinyo que sempre em dona.

A la Carla i l'Eli per ser una part imprescindible de qui sóc, a la Laia per totes les converses i per deixar-me sempre amb alguna reflexió interessant i a l'Anna perquè no ens haguem deixat de veure després de tants anys. Gràcies al Gil i al Jordi pels viatges i els no-viatges. A les Annes i a la Marina per tot el que hem viscut juntes, per cuidar-me i per ser la meva segona família aquests últims anys. Gràcies al Víctor, al Jaume, a la Glòria i a tots els de la uni, per aquests ja més de 10 anys junts, si no us hagués trobat no sé si estaria aquí ara mateix. I no puc deixar-me d'agrair a la JOC i a tots els que formeu part d'ella per permetre'm tenir un espai on desconnectar, gaudir de la música i poder seguir connectada amb Cerdanyola.

I finalment, moltíssimes gràcies a tota la meva família. Gràcies als meus iaios per tots els esforços que han fet al llarg de la seva vida, estic segura que d'altra forma el meu futur seria impensable. I especialment gràcies a les meves iaies per tota l'estima que em demostren constantment, ja sigui en forma d'abraçades i consells (no sempre ben rebuts) o en forma de churros, tomates, migas, picadillos y otras muchas comidas deliciosas. A la Denisse por sus consejos de edición y por mirarse esta tesis. I al Manel per deixar-me ser tal i com sóc i per poder-lo abraçar i molestar a parts iguals. I per últim, l'agraïment més necessari va dirigit als meus pares, per tota la seva dedicació i els seus esforços perquè jo pugui seguir endavant, per inculcar-me l'interès d'aprendre i de descobrir coses noves, per ajudar-me sempre i per cuidar-me, cada un a la seva manera. Gràcies per les teràpies abans i després de cada situació difícil i per la vostra confiança, clarament no hauria arribat fins aquí si no fos per vosaltres. Gràcies Immins, gràcies Mànol!

Contents

List of Symbols

Phylogenetics

$T; E(T); V(T)$	Tree, edges of the tree and vertices of the tree, page 17.
$\mathcal{T}_n$	Set of all isomorphic classes of unrooted trivalent n-leaf phylogenetic trees, page 20.
$T_{ab\|cd}$	Quartet tree with edge split $ab\|cd$, page 22.
$\{\pi; \{M_e\}_{e \in E(T)}\}$	Set of substitution parameters of the Markov process on a phylogenetic tree T, page 25.
JC69	Jukes Cantor model, page 27.
K80	Kimura 3-parameter model, page 29.
K81	Kimura 2-parameter model, page 29.
GM	General Markov model, page 31.
GTR	General time-reversible model, page 32.
ϕ_T^M	Polynomial map associated to tree T and model $\mathcal{M}$, page 47.
$\mathcal{V}_T^{\mathcal{M}}$	Smallest algebraic variety containing the image Im ϕ_T^M, page 50.
$\mathcal{V}_T^+$	Stochastic phylogenetic region associated to T, page 50.
φ_T	Polynomial map associated to the tree T and the JC69 model on Fourier coordinates, page 81.

PHYLOGENETIC RECONSTRUCTION METHODS

$F \sim \text{Mult}(N; p)$	Vector of relative frequencies obtained from N independent samples from a multinomial distribution p, page 63.
`Erik+2`(F)	Normalized scores output by `Erik+2`, page 67.
`SAQ`(F)	Normalized scores output by `SAQ`, page 125.
$\mathcal{J}(F)$	Scores output by the paralinear method, page 131.
`ASAQ`(F)	Normalized scores output by `ASAQ`, page 133.
`ML`	Maximum likelihood, page 64.
`NJ`	Neighbor-joining algorithm, page 65.
`MP`	Maximum Parsimony, page 66.
QP	Quartet Puzzling method, page 68.
WO	Weight Optimization method, page 68.
WIL	Willson's method, page 69.

MATRICES AND TENSORS

$\mathcal{M}_n(\mathbb{R})$	Set of square $n \times n$ real matrices, page 34.
$\mathcal{H}_n$	Hyperplane in $\mathbb{R}^n$ given by $x_1 + ... + x_n = 1$, page 38.
Δ^n	Standard simplex in $\mathbb{R}^n$, page 38.
$\mathbb{M}_n$	Set of $n \times n$ matrices with rows summing to one, page 38.
$\mathbb{M}_n^+$	Set of $n \times n$ stochastic matrices, page 38.
$\mathbb{M}_n^*$	Markov group: invertible matrices in $\mathbb{M}_n$, page 39.
$M \otimes N$	Kronecker product of two matrices, page 39.
$p \cdot M_1 \otimes \cdots \otimes M_n$	Markov action: group action of $\times^n \mathbb{M}_k^*$ on the space of tensors $\mathcal{W}^{\otimes n}$, page 39.
$p *_i M$	Markov action $p \cdot (Id \otimes \cdots \otimes Id \otimes \underbrace{M}_{i} \otimes Id \otimes \cdots \otimes Id)$, page 40.
$p_{...+...}$	i-th marginalization of p, page 39.
$Flat_{A\|B}$	Flattening matrix with respect to the split $A\|B$, page 41.
$N_{i,j}$	Double marginalization of a tensor p on the coordinates different from i and j, page 108.
$p_k^{A\|B}$	k-th $A\|B$ leaf-transformation on a tensor p, page 109.

DISTANCES

$d_{RF}(T_1, T_2)$	Robinson-Foulds distance between two phylogenetic trees, page 21.
$l(e)$	branch length of the edge e, page 33.
$d_{par}(s_1, s_2)$	Paralinear distance between DNA sequences, page 34.
$\delta_k(M)$	Distance from a matrix M to the variety of matrices with rank less than or equal to k, page 37.
$\delta_{sym}(M)$	Distance from a matrix M to the set of symmetric matrices, page 37.
$\delta_{psd}(M)$	Distance from a matrix M to the set of positive semidefinite matrices, page 37.
$psd(M)$	Closest positive semidefinite matrix to M, page 37.
$d(p, \mathcal{V})$	Euclidean distance from $p \in \mathbb{R}^n$ to an algebraic variety $\mathcal{V} \subset \mathbb{R}^n$, page 46.

INTRODUCTION

The term *phylogeny*, introduced by Haeckel (1866), denotes evolutionary relationships between different entities such as species, genes, or genomes. Under the assumption that quantitative similarities between these entities are related to their common evolutionary history, by comparing characteristics (either morphological characters or gene or protein sequences) of contemporary organisms, phylogeneticists try to infer their evolutionary history.

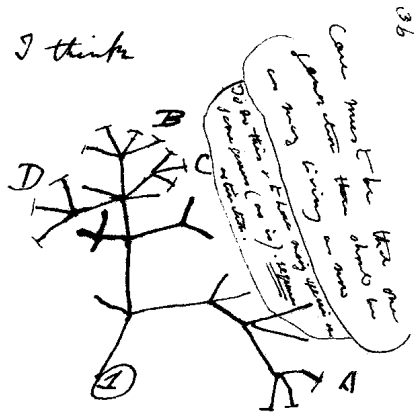


FIGURE 1: First diagram of an evolutionary tree by Charles Darwin on the *First notebook on transmutation of species, 1837*. The number 1 represents the ancestral species, letters adjacent to exterior edges are extant species and the remaining exterior edges represent extinct species.

The evolutionary relationships between entities are commonly depicted in the form of phylogenetic trees. However, phylogenetic trees go beyond evolutionary biology: other historical sciences such as linguistic history, which try to infer the history and relationship of languages that are currently spoken or have been spoken at some point, also make use of them. For these disciplines, trees are a basic tool to represent the elapsed time and the relationship between different events. The beginning of the use of trees is usually coined to Darwin for his early tree sketch of evolution (see Figure 1, 1837). However,

trees appear previously in Lamarck (1809) where the author sketched a diagram of animal groups with branching evolutionary paths, in the branching table of Germanic languages by Schottelius (1995) or in the work by Conrad *Gesner's Historiæ Animalium* (1555), in which a tree is drawn and used to classify species. Other disciplines such as comparative mythology or archaeology have recently been using phylogenetic and tree analysis in their studies (Renard, 2018).

In any case, where phylogenetic analysis is most used and represents the main challenge is in the field of evolutionary biology. In this setting, phylogenetics is the study of the evolutionary history and relationships among groups of biological entities (for instance species or genes) called *taxa*. According to the theory of the evolution of species developed by Darwin, species evolve through the natural selection of small variations that increase the individual's ability to compete, survive, and reproduce. These speciation processes are modeled by phylogenetic trees whose nodes represent different taxa and whose branches correspond to the evolutionary process between two taxa. The leaves of the tree are contemporary taxa and the root of the tree (if present) is the common ancestor to all the taxa represented on the tree. Throughout this work we consider phylogenetic trees whose leaves are labeled in correspondence to a set of taxa, see Figure 2.

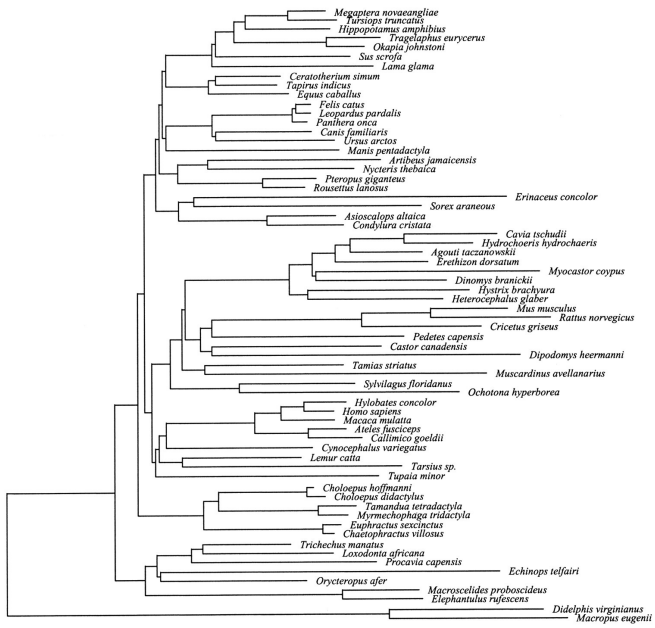


FIGURE 2: Phylogenetic tree of 64 placental mammals and two marsupials obtained by Murphy et al. (2001). Figure from Rosenberg and Kumar (2001).

During many years, phylogenetic trees were constructed based on comparisons among individuals of morphological or physiological characters. This was revolutionized in the 1960s and 70s with the advent of molecular data, first with protein sequences, later with nucleotide sequences and finally with the reconstruction of the entire genome. When such data became available, it was clear that similarities and differences in homologous nucleotide or protein sequences should be good indicators of relationships among taxa (Zuckerkandl and Pauling, 1965). Furthermore, biologists soon realized that genomic data could be used to infer phylogenetic relationships even among organisms that were so distantly related that they did not share apparent morphological or physiological characters.

The importance of phylogenetic analysis applies not only to the study of evolutionary relationships among organisms that host genes, but also to the study of genes themselves, making phylogenetics a useful tool in fields such as genomics (searching for genes), biomedicine (e.g. to trace cancer cells), and in the study of the origin of viruses. The most recent example of this application of phylogenetics is the detection of the origin of the SARS-CoV-2 virus, see Li et al. (2020).

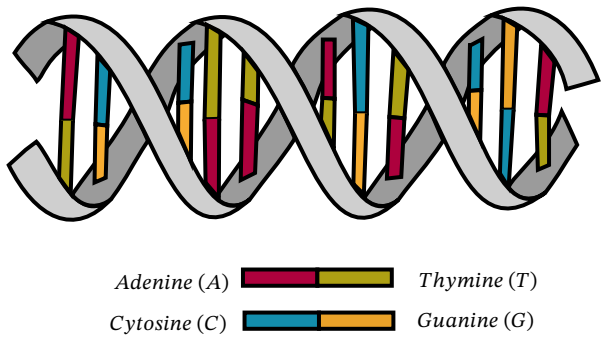


FIGURE 3: Illustration of a DNA molecule.

The genetic information of each individual is encoded in the *deoxyribonucleic acid* (briefly *DNA*) found at the nucleus of its cells. The DNA molecule consists of two strands of nucleotides which are coiled around each other to form a double helix. Each nucleotide is composed of a phosphate, a sugar and a basis. According to the bases, nucleotides are adenine (`A`), cytosine (`C`), guanine (`G`) and thymine (`T`). The nucleotides of the two strands match together to form a *base-pair* according to the Watson–Crick pairing $\texttt{A} - \texttt{T}$ or $\texttt{C} - \texttt{G}$ (see Cleaves (2011) and Figure 3 for an illustration). Due to this symmetry we can represent a DNA molecule as an ordered sequence of $\texttt{A}'s$, $\texttt{C}'s$, $\texttt{G}'s$ and $\texttt{T}'s$ (for example `ACCTGTATTC`), corresponding to one of the strands of the DNA.

The hereditary information from one generation to the next is supposed to be contained in the genes. The evolutionary process introduces changes in the genome, even in genetic sequences, that may complicate the comparison of DNA sequences. These changes in DNA sequences may be given by substitutions, insertions or deletions of nucleotides. Therefore sequences of the same gene at different species (or even at different individuals) may look very different. For that reason, the first problem in phylogenetic reconstruction is identifying which part of the DNA sequences of different taxa can be compared. That is, one needs to identify regions of similarity that may be a consequence of functional, structural, or evolutionary relationships between the sequences. Once this is identified, the information is collected in an *alignment*. We can represent an alignment by a table whose rows are the DNA sequences of the taxa considered and whose columns correspond to sites that (presumably) have evolved from the same site of a common ancestor (see Figure 4). Alignments are used in many contexts, and in the phylogenetic framework they are the basic ingredient to reconstruct phylogenetic trees. In most commonly used evolutionary models, insertions and deletions are not considered, as incorporating them would highly increase the complexity of the model. Therefore, all rows of the alignments we shall deal with have the same length, are comprised of words on the four letters `A,C,G,T`, and contain no gaps (which would represent an insertion or deletion). We refer the reader to Gusfield (1997) for an introduction to the subject of sequence alignments.

Taxa 1	`AAGCTTACATCCACCGTCGTTCTCATAAGTACGGGTCTCCCCCCGC`
Taxa 2	`CCTAAGCTTTCACATAATCGCCACCGGCGGACTTTTATCCACGATC`
Taxa 3	`CGGCGGCTTCCTTATAATTGACTCCACGTTCACTACATCAGTAAGC`

FIGURE 4: A multiple sequence alignment of DNA sequences of three taxa.

Obtaining a reliable alignment for a given set of DNA sequences is a complex task. There are different alignment methods available (see for example `BLAST` (Altschul et al., 1990) or `MAVID` (Bray and Pachter, 2003)), but it is out of the scope of this work to explain them. Throughout this work, we assume biological data is given through an alignment of DNA sequences.

The main problem we deal with in this thesis is the problem of *phylogenetic reconstruction*. That is, we aim to estimate the phylogenetic tree that best explains the evolutionary relationships of current taxa using solely information from their genome (in the form of an alignment). There are several phylogenetic reconstruction methods available, for example we highlight the *Maximum Likelihood* method by Felsenstein (1973) (`ML` for short, see Section 1.4.1 for an introduction) and the *Neighbor-joining* algorithm by Saitou

and Nei (1987) (briefly NJ, see Section 1.4.2) among the most commonly used. We refer to Warnow (2017) for an introduction and overview of phylogenetic inference methods. We focus on the reconstruction of the *topology* of phylogenetic trees (i.e., the shape of the tree taking into account the labels at the leaves), but we do not attempt to infer substitution parameters. Although the original motivation of the problem is biological, in the last decades mathematicians, statisticians and computer scientists have been working on the topic and providing useful insight. The approach followed by mathematicians is prominently focused on modeling the substitutions of nucleotides.

To this end, one usually assumes that different sites in the DNA sequences evolve independently and identically distributed (briefly iid), and according to a Markov process on a phylogenetic tree ruled by a model of nucleotide substitutions. These substitution models are specified by 4×4 transition matrices M_e associated to the edges e of the tree and by a distribution π of nucleotides at the root. Then, given an n-leaf tree T one can compute the distribution of nucleotide patterns at the leaves of the tree (representing the current taxa) in terms of the parameters of the model. This joint distribution is usually represented as a vector $p \in \mathbb{R}^{4^n}$ whose entries can be expressed as polynomials on the model parameters (see Sections 1.1.2 and 1.1.3). To each tree T, we can associate a polynomial map $\phi_T : \mathbb{R}^d \to \mathbb{R}^{4^n}$ that sends any d-tuple of parameters to a distribution vector of the 4^n possible observations at the leaves:

$$\begin{array}{rcl} \phi_T : \mathbb{R}^d & \longrightarrow & \mathbb{R}^{4^n} \\ \left(\pi; \{M_e\}_{e \in E(T)}\right) & \longmapsto & p = (p_{\mathtt{A},\mathtt{A}\ldots,\mathtt{A}}, p_{\mathtt{A},\mathtt{A}\ldots,\mathtt{C}}, p_{\mathtt{A},\mathtt{A}\ldots,\mathtt{G}}, \ldots, p_{\mathtt{T},\mathtt{T}\ldots,\mathtt{T}}). \end{array}$$

The transition matrices M_e and the distribution π may be required to satisfy additional constraints under some specific *evolutionary model*. The most restrictive model is the *Jukes Cantor model* (*JC69* for short, see Definition 1.1.25), the least restrictive is the *general Markov model* (briefly *GM*, see Definition 1.1.30), when no restrictions are imposed, and the models in between include *Kimura 3-parameter model* (*K81* for short, see Definition 1.1.27) for instance. All these models are considered at some point in this memoir.

Throughout this work we assume that the alignment we are given as input for the reconstruction problem is obtained from N independent samples of a multinomial distribution with parameters $p = \phi_T\left(\pi; \{M_e\}_{e \in E(T)}\right)$. If $F \in \mathbb{R}^{4^n}$ is the vector of relative frequencies of nucleotide patterns in this alignment, then we write $F \sim \text{Mult}(N; p)$.

This general approach for modeling evolutionary processes allows independent substitution probabilities, and in particular, different substitution rates of nucleotides, at different lineages. Some classical reconstruction methods, such as ML, that perform well

when applied to models with few parameters, may present some limitations when this heterogeneity at different lineages is assumed (as the number of such parameters increases). However, assuming heterogeneity of the substitution process between lineages is essential when inferring relationships between distantly related taxa (see for instance Felsenstein, 1978; Yang, 1994; Yang and Roberts, 1995; Galtier and Gouy, 1998; Ho and Jermiin, 2004; Foster, 2004). As we use the aforesaid evolutionary models, all the results presented in this memoir admit this heterogeneity across lineages.

The approach above naturally motivates the use of algebraic tools for phylogenetic reconstruction purposes. Biologists Lake (1987), Cavender and Felsenstein (1987) presented non-trivial algebraic relationships that were satisfied by any distribution p arising on the tree T (i.e. in the image of ϕ_T). Moreover, these relationships were not satisfied by distributions arising on the other phylogenetic trees $T' \neq T$ relating the same set of taxa, so they were potentially useful as phylogenetic inference methods. These relationships are known as *topology invariants* as they reflect the *tree topology*. In the last two decades, mathematicians working in the fields of algebraic geometry, commutative algebra and combinatorics have developed further successful results in this direction by studying the geometry of the algebraic varieties defined by $\operatorname{Im}\phi_T$, called *phylogenetic varieties*. See for example Allman and Rhodes (2008); Casanellas and Fernández-Sánchez (2008); Allman, Degnan, and Rhodes (2013); Chifman and Kubatko (2015); Allman, Kubatko, and Rhodes (2016); Casanellas, Fernández-Sánchez, and Michałek (2017), among others.

When reconstructing the tree topology using algebraic tools, the main approaches that have been followed are either using topology invariants (Lake, 1987; Casanellas and Fernández-Sánchez, 2007) or using the rank conditions on the *flattening* matrices (see Definition 1.2.23) arising from a certain rearrangement of the entries of distribution p at the leaves. That is, if this flattening matrix is constructed according to a certain split $A|B$ of the leaves obtained by removing an internal edge of the tree, then such flattening matrix, denoted by $Flat_{A|B}(p)$, has always rank less than or equal to 4, see Theorem 1.3.24. These are the main ideas under the software `SVDquartets` (Chifman and Kubatko, 2014; Chifman and Kubatko, 2015), `Split Scores` (Allman, Kubatko, and Rhodes, 2016) and the inference method `Erik+2` (Fernández-Sánchez and Casanellas, 2016), that will be mentioned later, is also based on this result (see Section 1.4.4 for an introduction).

All these tools rely on the fact that the distributions p that have arisen from a tree satisfy certain *algebraic* constraints, but they ignore the fact that they are actually distributions which arise from *stochastic* matrices at the edges of the tree (i.e. with positive entries and rows summing to one). These extra conditions lead to *semi-algebraic* constraints

which have been specified for certain models: Allman, Rhodes, and Taylor (2014) characterized the general Markov model, Matsen (2009) dealt with the Kimura 3-parameter model and Zwiernik and Smith (2011); Klaere and Liebscher (2012) with the 2-state case (2×2 transition matrices). This gives rise to the first question that motivated the research on this thesis:

Problem 1. Could semi-algebraic tools add some insight to the already existent algebraic tools for phylogenetic reconstruction?

Combining algebraic and semi-algebraic conditions to develop a tool for reconstructing the tree topology is not an easy task and, as far as we are aware, Kosta and Kubjas (2019) are the only ones that have used both tools together for the simple case of 2 states.

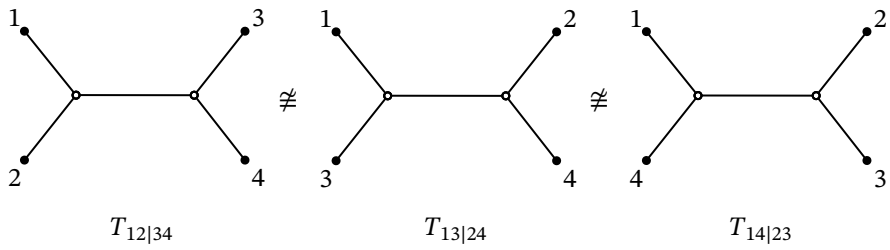


FIGURE 5: The three unrooted trivalent phylogenetic trees on 4 leaves: 12|34 (left), 13|24 (middle) and 14|23 (right).

We consider trees on four taxa, called 1, 2, 3 and 4 for simplicity. There are three unrooted and trivalent 4-leaf phylogenetic trees (or *quartets*), $T_{12|34}$, $T_{13|24}$, and $T_{14|23}$, depending on how the leaves are divided after removing the internal edge (see Figure 5 for an illustration of such quartets). We focus on quartets, as they may be considered as the basic ingredient to reconstruct larger trees. Indeed, there exist the so-called *quartet-based methods*, which infer the topology of larger trees T based on the reconstruction of all quartets subtrees of T.

For a given nucleotide substitution model, the Markov process on a quartet tree T gives rise to an algebraic variety $\mathcal{V}_T = \overline{\mathrm{Im}\phi_T}$ (by taking the Zariski closure). It is well known that the three *phylogenetic varieties* $\mathcal{V}_{T_{12|34}}$, $\mathcal{V}_{T_{13|24}}$, $\mathcal{V}_{T_{14|23}}$ do not coincide, which allows translating the topology reconstruction problem into the following problem: for a given distribution $p \in \mathbb{R}^{4^4}$ (which can be estimated from an alignment of four sequences as the observed relative frequencies of nucleotide patterns), decide to which of these three varieties p is closest (in terms of a certain distance or another specified optimization problem such as likelihood maximization). However, if we take into account that p is assumed to be obtained from *stochastic parameters* on one of these trees, then one

only should consider the *stochastic phylogenetic regions* of these varieties, $\mathcal{V}^+_{T_{12|34}}$, $\mathcal{V}^+_{T_{13|24}}$, $\mathcal{V}^+_{T_{14|23}}$, consisting on those points that are image of stochastic parameters.

Problem 2. Do semi-algebraic conditions support the same tree T whose algebraic variety $\mathcal{V}_T$ is closest to the data point?

In terms of distances, we explicitly ask the question:

$(*)$ If $p \in \mathbb{R}^{4^4}$ is a distribution satisfying $d\left(p, \mathcal{V}_{T_{12|34}}\right) < \min\left\{d\left(p, \mathcal{V}_{T_{13|24}}\right), d\left(p, \mathcal{V}_{T_{14|23}}\right)\right\}$, would it be possible that $d\left(p, \mathcal{V}^+_{T_{12|34}}\right) > \min\left\{d\left(p, \mathcal{V}^+_{T_{13|24}}\right), d\left(p, \mathcal{V}^+_{T_{14|23}}\right)\right\}$?

Figure 6 illustrates the hypothetical situation posed in this question.

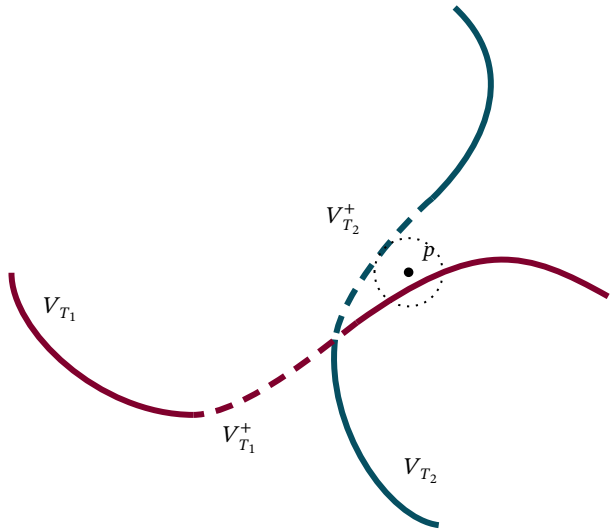


FIGURE 6: Illustration for the situation presented in Problem 2. The two curves represent two different phylogenetic varieties and their dashed parts corresponds to their stochastic phylogenetic regions. In this diagram p is closer to $\mathcal{V}_{T_1}$ than to $\mathcal{V}_{T_2}$. However, if we restrict to the stochastic regions, it is closer to $\mathcal{V}^+_{T_2}$ than to $\mathcal{V}^+_{T_1}$.

In Chapter 2 we address Problem 2 for special cases of interest in phylogenetics. First in Section 2.2 we consider trees with short branch lengths[1] at the external edges with nucleotides evolving under the Kimura 3-parameter model. Our main theoretical result in this sense is:

[1]Throughout this work, the *length of the branches* of phylogenetic trees refer to the amount of elapsed substitutions per site between both sequences at the ends of the edge.

Theorem 1 (See Theorem 2.2.3 for more precise statement). *Let $p_0 = \phi_T(Id, Id, Id, Id, M)$ where M is a $K81$ non-stochastic matrix and Id stands for the identity matrix. If $T' \neq T$ and p is close enough to p_0, then*

$$d(p, \mathcal{V}_T^+) < d(p, \mathcal{V}_{T'}^+).$$

Therefore we find that in this setting, restricting to the stochastic phylogenetic region does not make any difference and then the question posed in Problem 2 has a positive answer. It is worth noting that most reconstruction methods are successful on data points p close to a point p_0 as in the statement of the theorem.

The situation completely changes when we consider a long branch attraction scenario. The *long branch attraction* problem, *LBA* for short, occurs when fast evolving lineages are wrongly inferred to be closely related. Quartet trees representing these events are characterized for having two non-sister taxa that have accumulated many substitutions and two non-sister taxa that have very similar DNA sequences, see Figure 1.11 for an example. This is one of the most difficult situations to cope with in phylogenetic inference (see Kück et al., 2012).

As shown in the next result for the Jukes-Cantor model, considering the stochastic region of phylogenetic varieties might be of interest when dealing with data subject to LBA, specially if the data points are close to the intersection of the three varieties:

Theorem 2 (See Theorem 2.4.10). *Let $T = T_{12|34}$ and $p_0 := \phi_T(K_0, Id, K_0, Id, M_0) \in \mathcal{V}_T$ with K_0 and M_0 two JC69 matrices such that K_0 is stochastic and M_0 is not. If p is close enough to p_0 and p_0 satisfies a technical assumption (see Theorem 2.4.10), then for $T' \neq T$*

$$d(p, \mathcal{V}_T^+) \geq d(p, \mathcal{V}_{T'}^+).$$

We use techniques from *elimination theory* to prove this result and software from *numerical algebraic geometry* to do some simulations in this context, see Section 2.5. Theorem 2 together with the results of our simulations, imply that the answer to the question posed on Problem 2 is negative for data affected by the long branch attraction.

Following the last two theorems and relying on the results obtained in the simulations of Section 2.5, we conclude that incorporating the semi-algebraic conditions to the problem of phylogenetic reconstruction may be of importance when data are close to the

intersection of the three phylogenetic varieties (this is the case where phylogenetic reconstruction methods tend to confuse the trees); on the contrary, on data points which are far from the intersection it does not seem necessary to incorporate these semi-algebraic tools.

This suggests a positive answer to the question of Problem 1 but also reveals why incorporating semi-algebraic tools into phylogenetic reconstruction methods might be extremely difficult and the procedure to do it is not at all evident. Theoretically, this could be done via the semi-algebraic description of the general Markov model presented in Allman, Rhodes, and Taylor (2014). Nevertheless, in the same way that algebraic methods do not directly use topology invariants, the semi-algebraic conditions cannot be used straightforward or separately from the algebraic ones. Indeed, designing a method that works well for theoretical data $p \in \operatorname{Im}\phi_t$ is equivalent to specifying a method that is successful on distributions $F \sim \text{Mult}(N; p)$ when N tends to infinite. Which is more important is to design a method that works for F even if N is small and that is based on theoretical results states for F (not only for theoretical distributions p). These considerations suggest our next problem:

Problem 3. Merge the algebraic and semi-algebraic description for the general Markov model into a phylogenetic reconstruction method that performs well even with short alignments.

Briefly speaking, the result of Allman, Rhodes, and Taylor (2014) can be stated as follows. Let $p \in \mathbb{R}^{4^4}$ be a distribution and consider the 12|34 *leaf-transformations* on p: these transformations produce new vectors p_k that have also arisen on $T_{12|34}$ but with different transition matrices at the exterior edges. Then, p has arisen under some stochastic parameters on $T = T_{12|34}$, that is $p \in \operatorname{Im}\phi_{T_{12|34}}$, if and only if the following three conditions are satisfied:

(a) the marginalization of p over each leaf comes from stochastic parameters on a trivalent 3-leaf tree,

(b) the matrix $Flat_{12|34}(p)$ has rank four (or less than four for special parameters),

(c) after applying the 12|34 leaf-transformations to p, the flattening matrices $Flat_{13|24}(p_k)$ and $Flat_{14|23}(p_k)$ are symmetric and positive definite (or positive semidefinite for special parameters).

The first condition (a) is independent of the tree topology, so it seems not useful for recovering the tree topology. Condition (b) relies already on the tree topology and, as previously mentioned, has been used in different phylogenetic reconstruction methods.

However, it does not reflect the stochastic conditions of the transition matrices; instead condition (c) does: it reflects the stochastic nature of the transition matrix at the interior edge (see the proof of Theorem 1.3.26).

In order to combine conditions (b) and (c) and apply them in a reconstruction method, we need to guarantee that the rank cannot increase when projecting a matrix to the space of positive semidefinite matrices as we need to apply this result to distributions F that are not in $\mathrm{Im}\phi_T$. This is the content of the following result (see Corollary 3.1.3).

Theorem 3. *The rank of the positive semidefinite approximation of a real square matrix M is less than or equal to the rank of M.*

In Chapter 3 we present two quartet reconstruction methods that combine conditions (b) and (c) and achieve a high performance on simulated data. First, we present the method `SAQ`, which stands for *Semi-Algebraic Quartet reconstruction method* in Section 3.3. It is a phylogenetic reconstruction method for DNA alignments on four taxa which assumes the general Markov model on iid sites, but it can also deal with certain mixture models, that allow different behaviour on different sites of the DNA sequences. If $T = T_{12|34}$, we consider the 12|34 leaf-transformations p_k of p and for each we compute

$$s^k_{12|34} := \frac{\min\left\{\delta_4\left(psd\left(Flat_{13|24}\left(p_k\right)\right)\right), \delta_4\left(psd\left(Flat_{14|23}\left(p_k\right)\right)\right)\right\}}{\delta_4\left(psd\left(Flat_{12|34}\left(p_k\right)\right)\right) + \delta_4\left(psd\left(Flat_{12|43}\left(p_k\right)\right)\right)},$$

where δ_4 indicates the distance to the set of matrices of rank less than or equal to 4 and $psd(M)$ is the closest positive semidefinite matrix to M. `SAQ` also computes these scores for the other two quartets $s^k_{13|24}$ and $s^k_{14|23}$ and computes the normalized three scores:

$$\texttt{SAQ}(p) := \frac{1}{s}\left(s_{12|34}(p), s_{13|24}(p), s_{14|23}(p)\right),$$

where $s_{ab|cd}(p) = \mathrm{mean}_k\left\{s^k_{ab|cd}\right\}$ and $s := s_{12|34}(p) + s_{13|24}(p) + s_{14|23}(p)$. Finally `SAQ` outputs as the correct topology for a distribution p the topology $T_{A|B}$ that maximizes the corresponding value $s_{A|B}(p)$.

Theorem 4 (See Theorem 3.3.1)**.** *If p is a distribution that has arisen on the quartet $T_{12|34}$ with generic stochastic parameters and F is the vector of relative frequencies of patterns obtained from N independent trials sampled from p, then*

$$\lim_{N\to\infty} \texttt{SAQ}(F) = \texttt{SAQ}(p) = (1, 0, 0).$$

Moreover, the topology reconstruction method that assigns to a distribution F the quartet $T_{A|B}$ such that $s_{A|B}(F)$ is maximum is a statistically consistent method for the general Markov model.

In our studies on the performance of SAQ as a quartet reconstruction method, we have observed that SAQ tends to outperform Erik+2 for short alignments (length ≤ 1000) and in contrast Erik+2 obtains better results for longer alignment lengths. This fact has been the motivation for us to develop the ASAQ method. ASAQ is presented in Section 3.5, it is also a topology reconstruction method for four taxa that assumes a general Markov model of nucleotide substitution and iid distributed sites. It relies on a statistic $I_{A|B}(p)$ (see Section 3.4) that allows ASAQ to benefit from the advantages of Erik+2 when there is enough data or to draw on SAQ when there is an inconsistency between the value of the statistic and the tree output by Erik+2. This statistic is a "neighborliness" measure which coincides with the measure in Gascuel (1994) and it can be seen as assessing the positivity of the estimated interior branch length via the paralinear distance (see Definition 1.1.33). If we have a distribution p arising on a quartet $T_{A|B}$ under some stochastic parameters, then the value of $I_{A|B}(p)$ is positive and only depends on the internal transition matrix and the distribution at the root; in fact, it corresponds to twice the length of the interior branch. In Theorem 3.4.5 we see that the triplet $\mathcal{I}(p) = (I_{12|34}(p), I_{13|24}(p), I_{14|23}(p))$ is a quartet-inference measure (in the sense of Sumner et al., 2017) which leads to the *paralinear method*: given the vector F of relative frequencies of patterns in an alignment, choose the tree $T = T_{A|B}$ with largest value $I_{A|B}(F)$.

The method ASAQ (whose name arises from *Algebraic and Semi-Algebraic Quartet reconstruction method*) arises by combining this last method with SAQ and Erik+2. It proceeds as follows: given a distribution $F \in \mathbb{R}^{4^4}$ we compute the $\mathcal{I}(F)$ and the weights of Erik+2 at F. Then

(1) if Erik+2 and the paralinear method output the same quartet, ASAQ outputs the topology and weights of Erik+2,

(2) if they do not agree, then ASAQ outputs the weights of SAQ.

An inconsistency between Erik+2 and the paralinear method indicates that the algebraic conditions used by Erik+2 are not in concordance with the fact that the substitution parameters at the interior edge must be *stochastic*. Thus relying on SAQ in this case is a good option. We prove that ASAQ is statistically consistent in Theorem 3.5.1.

Although in this work SAQ and ASAQ have been specifically designed to deal with nucleotide sequences, the theoretical results underlying them could be generalised to

any number of states, and thus these methods could be adapted to deal with amino acid substitution models, which in some cases might be useful.

Proposing a statistical consistent method on phylogenetic reconstruction is not enough. We have tested the proposed methods for phylogenetic reconstruction on simulated (and real) data in order to check their actual performance in a number of scenarios, both consistent and violating the assumptions of the method. To this end we address the following goal:

Problem 4. Implement the proposed methods and test them on simulated data for quartets. Also, implement the methods as input of quartet-based methods and test them on simulated data for larger trees.

We address this problem in Chapter 4, where we present the performance of `SAQ` (see Section 4.2) and `ASAQ` (Section 4.4) on simulated data under different settings: on a "tree space" of quartets, on quartets with random branch lengths, and on alignments that are not identically distributed across sites ("mixture data" that violates the assumptions underlying `SAQ`). Data have been generated both under the GM model and a homogeneous across lineages and *general time-reversible model* (*GTR*, see Definition 1.1.31).

Our results show that both methods `SAQ` and `ASAQ` are highly successful, even with short alignments and with data that violates their assumptions. The weights ouput by `SAQ` and `ASAQ` are shown to be not biased towards any incorrect topology. We also provide comparisons of both methods against existing methods such as Maximum Likelihood, Neighbor-joining and `Erik+2`. Our simulation studies show that, in general, `SAQ` and `ASAQ` outperform all of them for GM data and have a compatible performance for GTR data (see Table 1). Moreover, the results on mixtures of distributions on the same quartet topology show that these methods can also deal with heterogeneity across lineages, as it surpasses methods that are specially designed for these data.

Average success of different methods on the tree space for GM data.

base pairs	`ASAQ`	`SAQ`	`Erik+2`	`NJ`	`ML`
500	**85.3**	84.6	72.4	72.5	72.1
1 000	**90**	88.8	80.3	79.7	73.6
10 000	**98.4**	96.8	97.1	94.3	75.4

TABLE 1: Average of correctly reconstructed trees of `SAQ` and `ASAQ` corresponding to the tree space of Figure 4.1 (b), compared to the results of the performance of `Erik+2`, `NJ` and `ML`, see Section 4.1.1 for an explanation on the simulated data.

Lastly, in Section 4.5 we consider three quartet based methods (or Q-methods): quartet-puzzling (QP) by Strimmer and Haeseler (1996), the method proposed by Willson (1999) (WIL), and Weight Optimization (WO) by Ranwez and Gascuel (2001), which reconstruct n-leaf trees T by using weights from each quartet subtree of T (see Section 1.4.5 for a brief description of such methods). As pointed out by Ranwez and Gascuel (2001), the weaknesses of Q-methods are very likely due to the method of weighting the quartets rather than to the method of combining them. Moreover, the correct management of long-branch attraction is crucial for obtaining a successful quartet-based method (St. John et al., 2003). We compare the performance of those three Q-methods in combination with weighting systems obtained by `ASAQ`, `Erik+2`, `NJ` and `ML`. We use the simulation framework proposed by Ranwez and Gascuel (2001) but considering more general models of nucleotide substitution. We observe a huge improvement on the performance of Q-based methods when weights from `ASAQ` or `Erik+2` are considered. The highest success is obtained by WO with the weighting system of `ASAQ`. The success of this method is compatible with a global `NJ` and outperforms it in the presence of mixtures. In Figure 7 we present the performance of WO with the weighting systems of `ASAQ`, `Erik+2`, `ML` and `NJ` for data simulated on a tree of 12 leaves and the GM model (see Section 4.1.2 for a detailed description of the data).

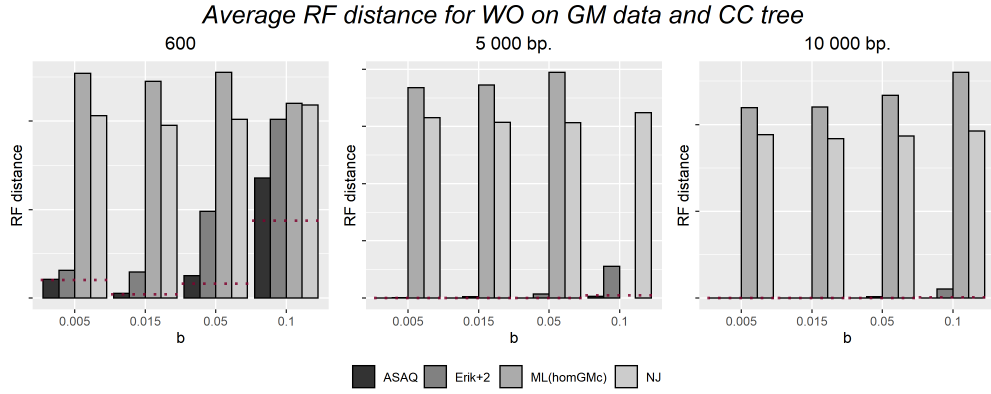


FIGURE 7: Average Robinson-Foulds (RF) distance to tree *CC* for WO with different weight systems (see Section 4.1.2), with data generated on CC trees with alignment length 600 bp. (left), 5 000 bp. (middle) and 10 000 bp. (right). The dotted red lines represent the average RF distance from the original tree to the tree reconstructed using a *global* `NJ` with paralinear distance.

Moreover, we considered real data for eight species of yeast provided by Jayaswal et al. (2014). WO with the weighting system of `ASAQ` correctly reconstructs the tree topology identified by Rokas et al. (2003) and widely accepted by the community of biologists.

As a final comment to this introduction, it is worth mentioning that the design of the methods SAQ and ASAQ is the result of multiple trials and simulations. As we have previously mentioned, how to combine algebraic and semi-algebraic conditions together is not obvious and it has taken several attempts to find an optimal and powerful way to do it. We have succeeded in proposing two statistically consistent methods that perform well both as quartet inference methods and as providing useful weighting systems to be used as input for Q-methods.

The structure of this thesis is as follows. In Chapter 1 we recall basic definitions, topics and results that will be needed later. We introduce phylogenetic trees and nucleotide substitution models and explain how to define a Markov process on a tree. Next, we give some background in tensors and algebraic geometry and describe phylogenetic algebraic varieties. We also introduce a characterization of the phylogenetic stochastic regions. Finally, we briefly describe existing phylogenetic reconstruction methods that will be used in the last chapter for comparison purposes. Chapter 2 is devoted to answering Problem 2 for trees with short branches at the external edges and trees under the long branch attraction phenomenon. Moreover, we present an algorithm to find the closest point in the stochastic phylogenetic region to a given data point and we provide some simulations to illustrate the positive answer for Question 2 on data close to the intersection of varieties. In Chapter 3 we present SAQ and ASAQ and prove the necessary technical results to explain their theoretical foundations. Lastly, in Chapter 4 we analyse the performance of SAQ and ASAQ as quartet inference methods on simulated data under different settings and we present the performance of different Q-methods using the weighted system provided by ASAQ in comparison to other methods, and test our methods with real data.

LIST OF PUBLICATIONS

Part of the results of this thesis appear in:

M. Casanellas, J. Fernández-Sánchez, M. Garrote-López, M. Sabaté-Vidales. **Phylogenetic reconstruction via quartet-based methods and semi-algebraic weights**, *In preparation* (2021).

M. Casanellas, J. Fernández-Sánchez, M. Garrote-López. **SAQ: semi-algebraic quartet reconstruction method**, *arXiv:2011.13968* (2020).

M. Casanellas, J. Fernández-Sánchez, M. Garrote-López. **Distance to the stochastic part of Phylogenetic Varieties**, *Journal of Symbolic Computation, Vol. 104, 563-682* (2021).

M. Garrote-López. **Computing the distance to the stochastic part of a phylogenetic variety.** *To appear in "Research Perspectives CRM Barcelona" Trends in Mathematics, Vol. 12, Springer-Birkhäuser* (2020).

M. Casanellas, J. Fernández-Sánchez, M. Garrote-López. **The inertia of the symmetric approximation of low rank matrices**, *Linear and Multilinear Algebra, Vol. 66(11), 2349-2353* (2018).

1

PRELIMINARIES

This chapter aims to introduce the preliminary concepts and results needed for the development of our work. The chapter is divided into four sections. In the first section, we introduce phylogenetic trees, Markov processes on a tree and nucleotide substitution models. Section 1.2 is devoted to introducing the basic notions and main tools on algebra and geometry needed on the sequel. In Section 1.3 we present the notions of phylogenetic algebraic varieties and their stochastic regions. Finally, in the last section, we explain what is the phylogenetic reconstruction problem we wish to address and give an overview on some existing phylogenetic reconstruction methods.

1.1 PHYLOGENETICS

1.1.1 Phylogenetic trees

A basic object of study in phylogenetics is a tree that represents the evolutionary relationships among a given set of taxa. In this section we introduce phylogenetic trees and some related notions. We follow the approach in the books by Allman and Rhodes (2003a), Allman and Rhodes (2005) and Steel (2016).

Definition 1.1.1. A *tree* T is a connected graph with no cycles. The set of vertices of T is denoted by $V(T)$ and the *degree* of a vertex is the number of edges incident on it. The vertices of degree 1 are called *leaves* and the set of leaves is denoted by $L(T)$. All other vertices, which have degree at least 2, are the *interior nodes* of the tree and the corresponding set is designated by $Int(T)$. $E(T)$ is the set of the edges of the tree and we call *exterior edges* those adjacent to a leaf of the tree. The remaining edges are called

interior edges. We shall write $e = \{u, v\}$ to indicate that the vertices u, v are the ends of the edge e.

Definition 1.1.2. A tree is called a *rooted tree* if one interior vertex r has been labeled as the "root". In this case, the edges of the tree are oriented away from the root.

Definition 1.1.3. Two trees T_1 and T_2 are *isomorphic* if there exists a graph morphism $\Psi : T_1 \to T_2$ consisting in a pair of bijective maps $\Psi_V : V(T_1) \to V(T_2)$ and $\Psi_E : E(T_1) \to E(T_2)$ such that if $e = \{u, v\} \in E(T_1)$ then $\Psi_E(e) = \{\Psi_V(u), \Psi_V(v)\} \in E(T_2)$. If T_1 and T_2 are rooted trees with respective roots r_1 and r_2 then, we also require that $\Psi_V(r_1) = r_2$.

Definition 1.1.4. For any two vertices $u, v \in V(T)$, if the path from u to v is oriented we say that u is an *ancestor* of v, and that v is a *descendant* of u. Moreover, if $e = \{u, v\}$ is an oriented edge $e : u \to v$, then we say that u is the *parent* of v, and v is the child of u (sometimes we shall also say that u and v are, respectively, the *parent node* and the *child node* of e). If a node has exactly two child nodes $u, v \in L(T)$ we say that (u, v) is a *cherry* of T.

Definition 1.1.5. A rooted tree is said to be *binary* if the root r has degree 2 and the other interior vertices are of degree 3. If T is an unrooted tree with all the interior vertices of degree 3 then it is called *trivalent*.

See Figure 1.1 for an example illustrating the last definition.

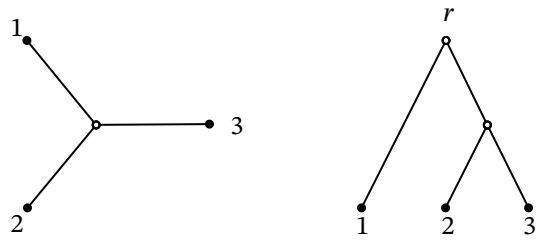


FIGURE 1.1: Left: A trivalent unrooted 3-leaf tree. Right: A rooted binary 3-leaf tree where (2, 3) is a cherry.

Definition 1.1.6. A *phylogenetic tree* is a pair (T, ψ) where T is a tree and $\psi : L(T) \to X$ is a one-to-one correspondence between the leaves of the tree and a finite set of labels (or taxa) X. We call ψ the *labeling map* of the phylogenetic tree.

In other words, the labeling map assigns a label to each leaf of the tree. In a phylogenetic tree, the set of labels X represents a set of usually non-extinct taxa, and the tree T shows the ancestral relationships among them. Every edge represents an evolutionary

process between the taxa represented by its end nodes. If the tree is rooted, the root represents the common ancestor of all the set of taxa X. For our purposes, X shall be usually taken as the set $[n] = \{1, 2, \dots, n\}$ and, since we identify the leaves of a phylogenetic tree with the set of labels X, we often use X when we refer to the set of leaves of T.

Definition 1.1.7. Two phylogenetic trees (T_1, ψ_1) and (T_2, ψ_2), with the same set of labels X at the leaves are *isomorphic* if there exists a tree isomorphism $\Psi : T_1 \to T_2$ such that for each leaf $u \in L(T_1)$, $\psi_1(u) = \psi_2(\Psi_V(u))$, where Ψ_V is the bijective map from $V(T_1)$ to $V(T_2)$. Roughly speaking, isomorphism of trees concern the shape of the trees and how the labels are placed at their leaves. We shall refer to this information as the *topology* of the tree. In terms of biology, two phylogenetic trees are isomorphic if they represent the same ancestral relationships among the set X (without taking into account time or evolutionary rates).

Example 1.1.8. Let T_1, T_2 be the trees of Figure 1.2. Consider a set of labels $X = [4]$ and the labeling maps:

$$\begin{array}{rcl@{\qquad}rcl} \psi_1 : L(T_1) & \to & X & \psi_2 : L(T_2) & \to & X \\ u_1 & \longmapsto & 1 & v_1 & \longmapsto & 3 \\ u_2 & \longmapsto & 2 & v_2 & \longmapsto & 4 \\ u_3 & \longmapsto & 3 & v_3 & \longmapsto & 2 \\ u_4 & \longmapsto & 4 & v_4 & \longmapsto & 1 \end{array} \tag{1.1}$$

Then, the phylogenetic trees (T_1, ψ_1) and (T_2, ψ_2) are isomorphic and its isomorphism class can be represented by the phylogenetic tree of Figure 1.3. If we root (T_1, ψ_1) at the vertex u_5 then it is isomorphic to (T_2, ψ_2) if and only if T_2 is rooted at the vertex v_6. On the other hand, if we consider the labeling map $\psi : L(T_2) \to X$ with $\psi(v_1) = 1$, $\psi(v_2) = 3$, $\psi(v_3) = 2$, and $\psi(v_4) = 4$, then (T_1, ψ_1) and (T_2, ψ) are not isomorphic.

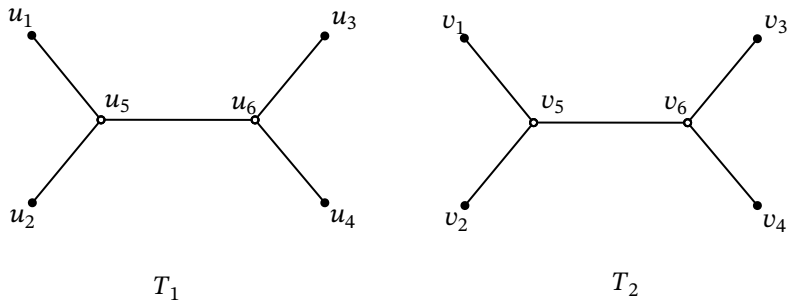


FIGURE 1.2: Two isomorphic trees T_1 and T_2.

From now on, we shall denote the phylogenetic tree (T, ψ) by T, if the labeling map ψ is clear from the context.

Another type of information that can be displayed on a phylogenetic tree is given by the *edge lengths* or *branch lengths*[1]. They usually represent the number of nucleotide changes per position that have occurred along the evolutionary process represented by that edge (see Section 1.1.4). By default, throughout this memoir, a phylogenetic tree T does not have this metric information and can be identified to any other tree isomorphic to T. Often we blur the distinction between a phylogenetic tree T and the tree topology of T, and call T to any phylogenetic tree isomorphic to T. Henceforth, we work with phylogenetic trees where the leaves are identified with some labels and the trees shall be displayed as in Figure 1.3.

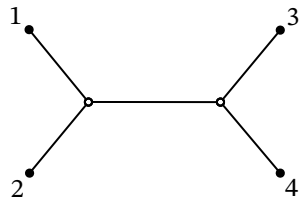


FIGURE 1.3: A 4-leaf phylogenetic tree isomorphic to (T_1, ψ_1) and (T_2, ψ_2) of Example 1.1.8.

Definition 1.1.9. We denote by $\mathcal{T}_n$ the set of all isomorphic classes of unrooted trivalent phylogenetic trees of n leaves.

Remark 1.1.10. It is well known that the number of trees in $\mathcal{T}_n$ for $n \geq 3$ is $|\mathcal{T}_n| = (2n-5)!!$ (see Lemma 2.33 in Pachter and Sturmfels, 2005).

Example 1.1.11. The first tree in Figure 1.1 is the only tree in $\mathcal{T}_3$, and will be denoted by T_3. The two trees (T_1, ψ_1) and (T_2, ψ) of the Example 1.1.8 are two different trees in $\mathcal{T}_4$.

Observe that the two isomorphic phylogenetic trees (T_1, ψ_1) and (T_2, ψ_2) of Example 1.1.8 have the same cherries $(1, 2)$, $(3, 4)$, and then, the way how the leaves are separated by the interior edge is the same. However, the phylogenetic tree (T_2, ψ) in the same example has two different cherries $(1, 3)$ and $(2, 4)$. This motivates our next definition:

Definition 1.1.12. A *split* $A|B$ of the taxa X of a tree T is a partition of the leaves into two non-empty subsets $A, B \subset X$ such that $A \cup B = X$, $A \cap B = \emptyset$. A split $A|B$ is *induced by the tree* (also called an *edge split*) if it is obtained by removing an edge e of T. We refer to splits not induced by a tree as *incompatible splits* of T. A split is called *trivial* if either A or B contain a single leaf.

[1] Throughout this work we will use edge and branch indistinguishably.

It is necessary to have a way to measure how similar different trees actually are. There are different ways to compare the structure or topology of phylogenetic trees, but in this work we will use the *Robinson-Foulds distance*, introduced by Robinson and Foulds (1981), which is defined as follows:

Definition 1.1.13. The *Robinson-Foulds distance* (RF distance for short) between two phylogenetic trees T_1 and T_2 with the same leaves X is the number of splits in one tree but not in the other one, and can be computed as

$$d_{RF}(T_1, T_2) = |\bar{E}(T_1)| + |\bar{E}(T_2)| - 2|S(T_1) \cap S(T_2)|, \tag{1.2}$$

where $\bar{E}(T_i)$ denotes the set of interior edges and $S(T_i)$ the set of non-trivial splits of T_i.

The RF distance is widely used by biologists, because it is fast to compute (it can be done in polynomial-time, see Steel, 2016) and it has a clear and biologically relevant interpretation: it is the number of evolutionary processes that take apart one tree from the other one. It can also be thought as the smallest number of edges that need to be collapsed in both trees to get the same unrooted phylogenetic tree.

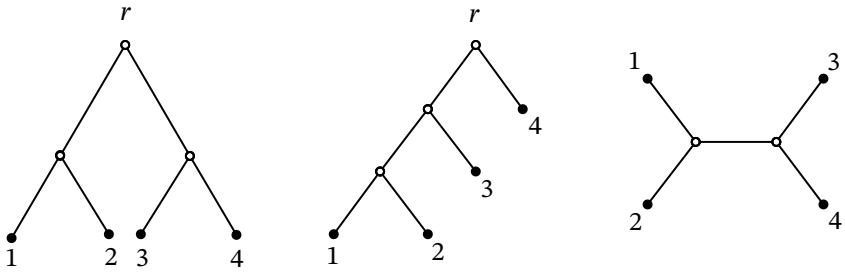


FIGURE 1.4: Three 4-leaf trees. The first two, are rooted and binary trees while the third one is unrooted and trivalent.

Example 1.1.14. Denote by T_1, T_2 and T_3 the phylogenetic trees at the left, middle and right (respectively) of Figure 1.4. Then $d_{RF}(T_1, T_2) = 2 + 2 - 2 \cdot 1 = 2$ and $d_{RF}(T_1, T_3) = 2 + 1 - 2 \cdot 1 = 1$. This agrees with the last interpretation, since we need to collapse an edge from T_1 and an edge from T_2 (2 edges in total) so that T_1 and T_2 are isomorphic unrooted trees. However, there is only one edge to be collapsed in T_1 to make it isomorphic to T_3.

Consider now the trees of Example 1.1.8 $T_1 := (T_1, \psi_1)$, $T_2 := (T_2, \psi_2)$ and $T_3 := (T_2, \psi)$. Then, $d_{RF}(T_1, T_2) = 1 + 1 - 2 = 0$ since they are isomorphic and $d_{RF}(T_1, T_2) = 1 + 1 - 0 = 2$ since we have to collapse the interior edge from both T_1 and T_3 to make them isomorphic.

Remark 1.1.15. Buneman (1971) proved that any two unrooted trivalent trees with the same set of edge splits are isomorphic. And as a consequence, two trees T_1 and T_2 with the same leaves X have the same topology if and only if $d_{RF}(T_1, T_2) = 0$.

Quartets

The 4-leaf trees, also called *quartets*, will play an important role in this work.

Definition 1.1.16. A *quartet* is an unrooted trivalent phylogenetic tree with 4 leaves. We denote a quartet by $T_{ab|cd}$ if $\{a, b\} \,|\, \{c, d\}$ is the non-trivial split induced by T.

Any phylogenetic tree induces a set of quartets. Namely, a quartet subtree $T_{aa'|bb'}$ is induced by a tree T if there exists an edge split $A|B$ of T such that $\{a, a'\} \subset A$ and $\{b, b'\} \subset B$.

Lemma 1.1.17 (Lemma 4.1, Steel, 2016)**.** *Let T_1 and T_2 be two isomorphic phylogenetic trees, then the collection of quartets induced by T_1 and T_2 are equal.*

Theorem 1.1.18 (Theorem 12, Allman and Rhodes, 2005)**.** *The collection of quartets induced by a trivalent tree T determines the topology of T .*

There are 3 different trivalent quartets that we denote $\mathcal{T}_4 = \{T_{12|34}, T_{13|24}, T_{14|23}\}$. These trees are presented in Figure 1.5.

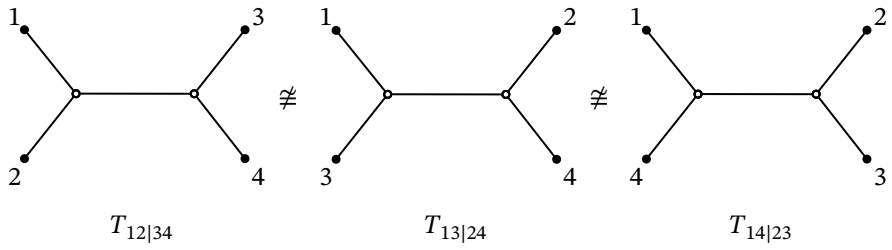


FIGURE 1.5: The three topologies of T_4: $T_{12|34}$, $T_{13|24}$ and $T_{14|23}$.

From now on we will always deal with phylogenetic trees, and we may use the terms tree or phylogenetic tree indistinctly.

1.1.2 Markov process on a tree

The substitution of nucleotides in an evolutionary process is usually modeled by adopting a parametric statistical model. In this sense the substitution of nucleotides in DNA sequences is assumed to be a random stochastic process. In order to model it we assume that different sites in the DNA sequence evolve independently, that is, the states

at each position in the sequence evolve independently of the other positions. Because of this hypothesis we can model substitutions in DNA sequences by modeling only one site. Moreover, for the moment we also assume that all sites in a DNA sequence evolve following the same process but we shall relax this hypothesis when we talk about mixture models (which are introduced at the end of this section).

Assuming the previous hypotheses, one models the substitution of nucleotides on a phylogenetic tree T as a Markov process on T. To this end, to each node u of T we associate a discrete random variable X_u; such X_u can take four different states on the set $\Sigma := \{\mathtt{A}, \mathtt{C}, \mathtt{G}, \mathtt{T}\}$ corresponding to the four nucleotides in DNA (denoted by their first letter). This can be generalized and assume each X_u takes k different states (for any natural k), however for our purposes it is enough to consider the set of nucleotides.

Before introducing the parameters of a model in a rooted phylogenetic tree T we introduce the notion of *stochastic vectors and matrices*:

Definition 1.1.19. A vector $v \in \mathbb{R}^n$ is a *stochastic vector* if all its entries are non-negative and sum to one. Sometimes, we will also refer to them as *distributions* since they can represent the probabilities that a random variable takes some prescribed n states. A *(row) stochastic matrix* M is a real square matrix with non-negative real entries and with rows summing to one. Stochastic matrices are also called *transition*, *probability*, *substitution* or *Markov* matrices.

Throughout this section we assume that trees are rooted. The parameters of a Markov process on T include the *distribution at the root* i.e. the distribution $\pi = (\pi_{\mathtt{A}}, \pi_{\mathtt{C}}, \pi_{\mathtt{G}}, \pi_{\mathtt{T}})^t$ of the random variable X_r associated to the root r. The entries of π are interpreted as the probabilities that an arbitrary site in the DNA sequence at the root is occupied by the corresponding nucleotide, or, equivalently, the nucleotide frequencies that we would expect to observe in a sequence at the root.

A second set of parameters is associated to the substitution processes that occur on the edges of the tree. To each edge $e \in E(T)$ we attach a 4×4 *transition* matrix M_e. The entries of M_e are the conditional probabilities $\Pr(x|y, e)$ that the state y at the parent node of e is being substituted by the state x at its child. In other words, the (i, j)-entry of M_e stands for the conditional probability that the nucleotide i at the DNA sequence at the parent vertex of the edge e, is replaced by nucleotide j at the descendant vertex.

In this case, the transition matrices have the form

$$M_e = \begin{array}{c} \\ \mathtt{A} \\ \mathtt{C} \\ \mathtt{G} \\ \mathtt{T} \end{array} \begin{array}{c} \begin{array}{cccc} \mathtt{A} & \mathtt{C} & \mathtt{G} & \mathtt{T} \end{array} \\ \begin{pmatrix} \Pr(\mathtt{A}|\mathtt{A},e) & \Pr(\mathtt{C}|\mathtt{A},e) & \Pr(\mathtt{G}|\mathtt{A},e) & \Pr(\mathtt{T}|\mathtt{A},e) \\ \Pr(\mathtt{A}|\mathtt{C},e) & \Pr(\mathtt{C}|\mathtt{C},e) & \Pr(\mathtt{G}|\mathtt{C},e) & \Pr(\mathtt{T}|\mathtt{C},e) \\ \Pr(\mathtt{A}|\mathtt{G},e) & \Pr(\mathtt{C}|\mathtt{G},e) & \Pr(\mathtt{G}|\mathtt{G},e) & \Pr(\mathtt{T}|\mathtt{G},e) \\ \Pr(\mathtt{A}|\mathtt{T},e) & \Pr(\mathtt{C}|\mathtt{T},e) & \Pr(\mathtt{G}|\mathtt{T},e) & \Pr(\mathtt{T}|\mathtt{T},e) \end{pmatrix} \end{array}.$$

It is easy to that any 4×4 stochastic matrix M has an eigenvalue equal to one, see Theorem 1.2.15 on Section 1.2.1. Moreover, if M is positive, then $\lambda = 1$ is the only eigenvalue with $|\lambda| = 1$, it has algebraic multiplicity one and there exists a unique distribution $\overline{\pi}$ such that $\overline{\pi}^t M = \overline{\pi}^t$.

Definition 1.1.20. When M is positive, the distribution $\overline{\pi}$ such that $\overline{\pi}^t M = \overline{\pi}^t$ is called the *stationary distribution* of M.

Definition 1.1.21. A *Markov process on a tree* T is the parametric model that expresses the joint probability distribution of the random variables $\{X_u\}_{u \in V(T)}$ as a product of entries of the transition matrices associated to the edges of T and the distribution π at the root:

$$Pr\left(\{X_u = x_u\}_{u \in V(T)}\right) = \prod_{x_r} \prod_{e \in E(T)} M_e(x_u, x_v),$$

where e is the oriented edge $e = u \to v$.

This Markov process can also be understood as a condition of independence on the random variables X_u. That is, a *Markov process on a tree* T is a collection of variables $\{X_u\}_{u \in V(T)}$ such that for each edge $e \in E(T)$, $e : u \to v$, X_v is conditionally independent of any variable X_w at the non-descendant nodes of v (with $w \neq u$) given X_u. Moreover, it can also be regarded as an extension of a Markov chain, since any path on a tree can be seen as a Markov chain composed of the nodes along the path.

Since DNA sequences of the contemporary taxa are known, but we do not have any information about the ancestral taxa we deal with a *hidden Markov process*. We shall say that the random variables at the leaves of T are *observed* and the ones at the interior nodes are *hidden*. In what follows we describe how to compute the joint probability at the random variables at the leaves according to the described Markov process on T. The probability that the random variables $X_1, \dots, X_n$ at the leaves take the states $x_1, \dots, x_n$ respectively is denoted by

$$p_{x_1,\dots,x_n} = \Pr(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n).$$

Then, we define *joint distribution* $p \in \mathbb{R}^{4^n}$ at the leaves of a rooted phylogenetic tree T as the vector p whose entries are the joint probabilities $p_{x_1 \dots x_n}$,

$$p = (p_{x_1,\dots,x_n})_{x_1,\dots,x_n \in \Sigma}.$$

According to the Markov process on the tree and marginalizing over the random variables at the leaves, we can express $p_{x_1,\dots,x_n}$ in terms of the entries of the substitution matrices as follows:

$$p_{x_1,\dots,x_n} = \sum_{x_r \in \Sigma} \sum_{\substack{x_v \in \Sigma \\ v \in \mathrm{Int}(T)}} \pi_{x_r} \prod_{e \in E(T)} M_e\left(x_{a(e)}, x_{d(e)}\right), \tag{1.3}$$

where x_r is the state of the root, $x_{a(e)}$ is the state of the parent node of e, and $x_{d(e)}$ is the state of the child node of e. If e is a terminal edge ending at leaf i then $x_{d(e)} = x_i$. Note that every entry of p can be seen as a polynomial on the parameters of the model $\mathcal{M}$.

Notation 1.1.22. From now on, the set $\left\{\pi; \{M_e\}_{e \in E(T)}\right\}$ consisting on the transition matrices and the distribution at the root on a tree, will be called the *substitution parameters* of the Markov process on T. In addition, we say that p is a distribution that *has arisen on T with certain substitution parameters* if p satisfies (1.3).

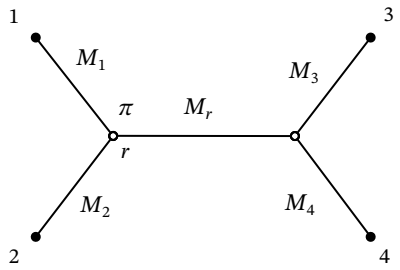


FIGURE 1.6: Markov process on a quartet.

Example 1.1.23. The Figure 1.6 represents a Markov process on a phylogenetic tree, where the X_i are random variables associated to the leaves, M_i are the transition matrices, and π is the distribution at the root r. We can compute the entry p_{x_1,x_2,x_3,x_4} of the joint distribution p using equation (1.3):

$$p_{x_1,x_2,x_3,x_4} = \sum_{x_r,x_5 \in \Sigma} \pi_{x_r} \cdot M_5(x_r, x_5) \cdot M_1(x_r, x_1) \cdot M_2(x_r, x_2) \cdot M_3(x_5, x_3) \cdot M_4(x_5, x_4). \tag{1.4}$$

A continuous-time Markov process on a tree

A Markov processes on a tree can also be described using a continuous notion of time. In this formulation, we suppose that there are certain *instantaneous rates* q_{xy} (with $x \neq y$) of substitutions from x to y which can be thought as the speed at which substitutions occur. We can set $q_{xx} = -\sum_{x,y} q_{xy}$ and arrange these rates into a *rate matrix*

$$Q = \begin{pmatrix} q_{AA} & q_{AC} & q_{AG} & q_{AT} \\ q_{CA} & q_{CC} & q_{CG} & q_{CT} \\ q_{GA} & q_{GC} & q_{GG} & q_{GT} \\ q_{TA} & q_{TC} & q_{TG} & q_{TT} \end{pmatrix},$$

that is, a matrix with non-negative real off-diagonal entries and rows summing to zero. Each entry $q_{xx} \leq 0$ is the rate of loss of x and each entry q_{xy} is the rate of changes from state x to state y. For that reason, the rates q_{xx} and q_{xy} must balance and the rows of the rate matrices must sum to zero.

Under this assumption, the matrix

$$M(t_e) = exp(Qt_e) \tag{1.5}$$

is a Markov matrix and encodes the conditional probabilities of substitutions at time t_e. A transition matrix M is called *embeddable* if there exists some rate matrix Q and value $t > 0$ such that $M = exp(Qt)$.

Usually a continuous-time Markov process on a phylogenetic tree is described by a rate matrix that remains constant during the whole process. In this case we say that the process is *homogeneous*.

Mixture models

We have assumed so far that all sites in a DNA sequence evolve identically and independently according to a given Markov process. However, this assumption is not realistic and for biological reasons sometimes one needs to relax it by assuming different behaviour on different sites of the DNA sequences. This can be modeled as a *mixture of Markov processes*.

Definition 1.1.24. Let T be a phylogenetic tree and let p and p' be two distributions that have arisen on T with substitution parameters $\left\{\pi; \{M_e\}_{e \in E(T)}\right\}$ and $\left\{\pi'; \{M'_e\}_{e \in E(T)}\right\}$,

respectively. Then, we say that a distribution q follows a *mixture model* with 2 categories on T if it satisfies

$$q = \lambda p + (1-\lambda)p',$$

where $0 \leq \lambda \leq 1$. In this case, λ represents the probability that a random site in the sequences at the leaves evolves under the Markov process associated with p.

Mixture models on phylogenetics can be generalized and a distribution can be assumed to be a convex combination of m different distributions $q = \sum_{i=1}^{m} \lambda_i p_i$ where $\sum_{i=1}^{m} \lambda_i = 1$ and each p_i has arisen on a tree T with different substitution parameters (see Steel, 2016). In this case q follows a mixture model with m categories.

1.1.3 Nucleotide substitution models

When one uses Markov processes to model the substitution of nucleotides on a DNA sequence, one can take into account certain restrictions (which may arise from biochemical properties of nucleotides) on the distribution at the root and the transition matrices. This is specified by a *nucleotide substitution model*.

We say that a phylogenetic tree T evolves under a certain nucleotide substitution model if the distribution at the root and the transition matrices attached to the edges of T are taken from that particular model. The resulting evolutionary model on T has a number d of free parameters, which can be bounded by $3 + 12|E(T)|$, where $|E(T)|$ is the number of edges of T, because the distribution at the root is assumed to sum to one (*i.e.* has at most 3 free parameters) and the rows of the transition matrices also sum to one (*i.e* there are no more than 12 free parameters in each matrix).

We present some nucleotide substitution models, starting with the most restricted case and then relaxing the conditions.

Jukes Cantor model

Definition 1.1.25. The *Jukes-Cantor model* (*JC69* for short, see the work by Jukes and Cantor, 1969) is the most restricted and at the same time, the simplest model. It assumes a uniform distribution at the root, $\pi = \left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right)^t$, equal probabilities of substitution and equal probabilities of no-substitution. Then, the transition matrices have the structure:

$$M_e = \begin{pmatrix} a_e & b_e & b_e & b_e \\ b_e & a_e & b_e & b_e \\ b_e & b_e & a_e & b_e \\ b_e & b_e & b_e & a_e \end{pmatrix}, \text{ where } a_e = 1 - 3b_e. \tag{1.6}$$

A *JC69 matrix M* is any matrix satisfying (1.6); if its entries are non-negative we say that M is a *stochastic JC69 matrix*. Note that we have one free parameter b_e per edge $e \in E(T)$ and the distribution at the root is fixed. The total number of free parameters of the JC69 model on a tree T is $|E(T)|$.

Example 1.1.26. Consider the Markov process on a phylogenetic tree presented in Figure 1.6 and assume it follows a JC69 model. Then, since π is fixed and each transition matrix has 1 free parameter, the number of free parameters in total is 6 (one for each matrix). We can compute the joint distribution p at the leaves and express each entry of p as a polynomial in terms of a_e and b_e (taking into account that $a_e = 1 - 3b_e$). For instance, if the same state $x \in \Sigma$ is observed at the four leaves, we have:

$$\begin{aligned} p_{x,x,x,x} &= \sum_{x_r,x_5\in\Sigma} \frac{1}{4} M_5(x_r,x_5)M_1(x_r,x)M_2(x_r,x)M_3(x_5,x)M_4(x_5,x) = \\ &= \frac{a5}{4}\left(a_1a_2a_3a_4 + 3b_1b_2b_3b_4\right) + \frac{3b_5}{4}\left(b_1b_2a_3a_4 + a_1a_2b_3b_4 + 2b_1b_2b_3b_4\right). \end{aligned}$$

However, if the observation at leaves 1 and 2 and 3 is x but the one at leaf 4 is y, with $x \neq y$, then the entries $p_{x,x,x,y}$ of p are

$$\begin{aligned} p_{x,x,x,y} = &\frac{a_5}{4}\left(b_1b_2b_3a_4 + a_1a_2a_3b_4 + 2b_1b_2b_3b_4\right) + \\ &\frac{b_5}{4}\left(a_1a_2b_3a_4 + 2b_1b_2b_3a_4 + 3b_1b_2a_3b_4 + 2a_1a_2b_3b_4 + 4b_1b_2b_3b_4\right). \end{aligned}$$

Kimura models

The JC69 model gives equal probabilities to each of the nucleotide substitutions. However, this is not completely realistic since substitutions are often more expected to occur within the *purine* group {A, G} and within the *pyrmidine* group {C, T}, than between different groups. The substitutions within the purine group or the pyrmidine group are called *transitions*, and substitutions from purine to pyrimidine or vice versa are called *transversions*. Kimura 1980, 1981 tried to adjust this by explicitly adding different parameters for transversions and transitions according to the diagram presented in Figure 1.7. The proposed models are called the *Kimura 3-parameter* model (K81 for short) if two different parameters are considered for the transversions, and the *Kimura 2-parameter* model (K80) if all transversions occur with the same probability.

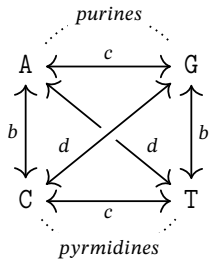


FIGURE 1.7: Horizontal arrows represent transitions while vertical and diagonal arrows are transversions.

Definition 1.1.27. The *Kimura 3-parameter model* (*K81* for short) assumes that the nucleotide distribution π at the root is uniform and that the transition matrices are of the form

$$M_e = \begin{pmatrix} a_e & b_e & c_e & d_e \\ b_e & a_e & d_e & c_e \\ c_e & d_e & a_e & b_e \\ d_e & c_e & b_e & a_e \end{pmatrix}, \tag{1.7}$$

where $a_e = 1 - b_e - c_e - d_e$. A *K81 matrix* M is any matrix of the form (1.7) with the rows summing to 1; if the entries of M are non-negative we say that M is a *stochastic K81 matrix*. In this model we have three free parameters per edge and the root distribution is fixed, so the number d of free parameters of a process on a tree T that follows this model is $d = 3|E(T)|$. In particular, for the phylogenetic tree in Figure 1.6, we have $d = 15$.

The more restricted *Kimura 2-parameter model* (Kimura, 1980) adds the additional assumption that $b_e = d_e$. Note that the JC69 model is a submodel of both Kimura models with $b_e = c_e = d_e$. For all these models $\pi = \left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right)$ is the stationary distribution of the transition matrices (see Definition 1.1.20).

One of the key interest of these models is the fact that all transition matrices diagonalize through the same basis of eigenvectors. This follows from the following easy lemma:

Lemma 1.1.28. *If M is a K81 matrix as in (1.7), then it diagonalizes with eigenvalues $m_{\mathtt{A}} = a_e + b_e + c_e + d_e = 1$, $m_{\mathtt{C}} = a_e + b_e - c_e - d_e$, $m_{\mathtt{G}} = a_e - b_e + c_e - d_e$ and $m_{\mathtt{T}} = a_e - b_e - c_e + d_e$ and respective eigenvectors $\overline{\mathtt{A}} = (1,1,1,1)^t$, $\overline{\mathtt{C}} = (1,1,-1,-1)^t$, $\overline{\mathtt{G}} = (1,-1,1,-1)^t$ and $\overline{\mathtt{T}} = (1,-1,-1,1)^t$. In other words, M can be written as*

$$M = H^{-1} \cdot \overline{M} \cdot H,$$

where $\overline{M}$ is the diagonal matrix with entries $m_{\mathtt{A}}, m_{\mathtt{C}}, m_{\mathtt{G}}, m_{\mathtt{T}}$ and the matrix of change of basis

$$H = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix} \tag{1.8}$$

is known as the Hadamard matrix. Moreover, $H^{-1} = \frac{1}{4}H^t = \frac{1}{4}H$.

In particular, the eigenvalues of a JC69 matrix are $m_{\mathtt{A}} = 1$ and $m_{\mathtt{C}} = m_{\mathtt{G}} = m_{\mathtt{T}} = 1 - 4b_e$.

Remark 1.1.29. Hendy, Penny, and Steel (1994) observed that there is a natural group action of $\mathbb{Z}_2 \times \mathbb{Z}_2$ on the set of nucleotides $\{\mathtt{A}, \mathtt{C}, \mathtt{G}, \mathtt{T}\}$ that acts transitively and freely and consistent with the Kimura K81 structure. This fact allows us to define the models above (JC69, K80 and K81 models) in terms of matrices that are invariant by this action. In this context, the Hadamard matrix presented above can be also understood as a discrete Fourier transform. The basis of eigenvectors denoted by $\overline{\Sigma} = \left\{\overline{\mathtt{A}}, \overline{\mathtt{C}}, \overline{\mathtt{G}}, \overline{\mathtt{T}}\right\}$ is called the *Fourier basis* and the eigenvalues $m_{\mathtt{A}}, m_{\mathtt{C}}, m_{\mathtt{G}}, m_{\mathtt{T}}$ of M_e, the *Fourier parameters*. In Section 1.3.2 we go deeper in the notion of *Fourier transform* and we introduce the *Fourier coordinates* of a distribution p.

In this setting JC69 and the Kimura models are instances of the so-called group-based models. A *group-based model* relative to an abelian group G is a substitution model whose matrices can be defined in terms of a function $f : G \to \mathbb{R}$ such that $M(g,h) = f(g-h)$, for any $g, h \in G$. If we consider the bijection between Σ and the group $G = (\mathbb{Z}_2 \times \mathbb{Z}_2, +)$ given by

$$\begin{array}{ccc} \Sigma = \{\mathtt{A}, \mathtt{C}, \mathtt{G}, \mathtt{T}\} & \longleftrightarrow & \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \\ \mathtt{A} & \longmapsto & (0,0) \\ \mathtt{C} & \longmapsto & (0,1) \\ \mathtt{G} & \longmapsto & (1,0) \\ \mathtt{T} & \longmapsto & (1,1) \end{array}$$

then the K81 model is a group-based model with a function f satisfying $f(0,0) = a_e$, $f(0,1) = b_e$, $f(1,0) = c_e$, $f(1,1) = d_e$, where a_e, b_e, c_e, d_e are the entries of a K81 matrix M_e as in (1.7). Note that with the identification of Σ with G we can define the sum operation on nucleotides, with $\mathtt{A}$ equal to the neutral element. For instance $\mathtt{C} + \mathtt{G} = \mathtt{T}$. The JC69 and K80 models are also group-based models with respect to $G = \mathbb{Z}_2 \times \mathbb{Z}_2$ since they are submodels of the K81 model. We refer to the work by Sturmfels and Sullivant (2005) for more details.

General Markov model

Definition 1.1.30. The *general Markov model* (briefly *GM* model) does not impose any restriction neither on the distribution π nor on the transition matrices M_e. Hence, $\pi = (\pi_{\mathtt{A}}, \pi_{\mathtt{C}}, \pi_{\mathtt{G}}, \pi_{\mathtt{T}})^t$ with $\sum_i \pi_i = 1$, and

$$M_e = \begin{pmatrix} a_e & b_e & c_e & d_e \\ e_e & f_e & g_e & h_e \\ j_e & k_e & l_e & m_e \\ n_e & o_e & p_e & q_e \end{pmatrix}, \text{ where } \begin{cases} a_e + b_e + c_e + d_e = 1, \\ e_e + f_e + g_e + h_e = 1, \\ j_e + k_e + l_e + m_e = 1, \\ n_e + o_e + p_e + q_e = 1. \end{cases} \tag{1.9}$$

A *GM matrix* M is any matrix satisfying (1.9); if the entries of M are non-negative we say that M *is a stochastic GM matrix*. The number of free parameters on a tree evolving under the GM model is $d = 3 + 12|E(T)|$ and for the particular case of the tree presented in Figure 1.6, $d = 12 \cdot 6 + 3 = 75$.

This model is the most general and will be important in the next chapters. It has been deeply studied in Allman and Rhodes (2004); Allman and Rhodes (2008); Draisma and Kuttler (2009) among others.

Time-reversible models

Consider an oriented edge e going from u to v. Let X_u and X_v be random variables associated to the vertices u and v and suppose π is the distribution at the ancestor u and M the transition matrix on e.

$$u \xrightarrow[\pi]{M} v$$

FIGURE 1.8: Markov process on an edge.

Let p be the joint distribution of X_u and X_v, with $p_{ij} = \Pr(X_u = i, X_v = j) = \Pr(X_u = i)\Pr(X_v = j | x_X = i) = \pi_i M(i,j)$. Thus, the joint distribution $p \in \mathbb{R}^{16}$ can be organized as the matrix $P = diag(\pi)M$, where $diag(\pi)$ is a diagonal matrix with the entries of π at the diagonal. We say that this process is *time-reversible* if $P = P^t$, or in other words, if

$$diag(\pi)M = M^t \mathrm{diag}(\pi).$$

When a model satisfies this equality we say it is *time-reversible*. All the models previously introduced except the GM model are time-reversible.

In the case of continuous-time models this condition translates to

$$diag(\pi)Q = Q^t diag(\pi).$$

The most general time-reversible model is known as the *general time-reversible* model (GTR for short) and was introduced by Tavaré (1986).

Definition 1.1.31. The *general time-reversible model* (*GTR*) is a continuous-time model specified by a distribution $\pi = (\pi_{\mathtt{A}}, \pi_{\mathtt{C}}, \pi_{\mathtt{G}}, \pi_{\mathtt{T}})^t$ at the root and a rate matrix with the form

$$Q = \begin{pmatrix} * & \pi_C a & \pi_G b & \pi_T c \\ \pi_A a & * & \pi_G d & \pi_T f \\ \pi_A b & \pi_C d & * & \pi_T g \\ \pi_A c & \pi_C f & \pi_G g & * \end{pmatrix}, \tag{1.10}$$

for some $a, b, c, d, f, g > 0$ and with sum of rows equal to zero. Each matrix of the form (1.10) is called a *GTR matrix*. When we refer to the GTR model, we assume homogeneity and so each matrix is $M_e = exp(Qt_e)$ for certain $t_e > 0$.

In the rest of this section we describe two particular mixture models (see Section 1.1.2). We consider a phylogenetic tree T with n leaves and we describe the mixtures of Markov processes on T.

Invariable sites models

A particular interesting class of mixture models are the *invariable sites models*. These models assume that some sites on the DNA sequence (the *variable* sites) evolve according to a certain model as described above and the rest of the sites are *invariable*, meaning that they do not change. If the variable sites follow a model $\mathcal{M}$ then, this mixture model is denoted as $\mathcal{M} + I$.

Example 1.1.32. If q is a distribution obtained from a model $GM + I$ then is satisfies

$$q = \lambda p + (1 - \lambda)p',$$

where $0 \leq \lambda \leq 1$, p has arisen from a general Markov model, and the entries of p' satisfy

$$p'_{x_1,\ldots,x_n} = \begin{cases} \pi_{x_1} & \text{if } x_1 = \cdots = x_n, \\ 0 & \text{otherwise,} \end{cases} \tag{1.11}$$

where $\pi = (\pi_{\mathtt{A}}, \pi_{\mathtt{C}}, \pi_{\mathtt{G}}, \pi_{\mathtt{T}})^t$ is the distribution at the root.

Gamma-distributed rates across sites

In this case we assume a GTR model where rates may vary across sites. One usually assumes that each site has its own rate r described by a gamma distribution Γ (or a discrete approximation of it) parametrized by a single parameter. This gives rise to the general time-reversible with *gamma-distributed rates model* GTR + Γ. For more flexibility, one can allow $r = 0$ with certain probability λ and assume r follows a gamma distribution with probability $1 - \lambda$, adding then a category of invariable sites. This is known as the *general time-reversible model with gamma-distributed rates and invariable sites model* GTR+Γ+I.

1.1.4 Evolutionary distances

A length assigned to an edge of a phylogenetic tree might have multiple interpretations, but it measures the difference between both sequences at the ends of the edge in some way. In this work we adopt the most common approach: the length of an edge, or *branch length*, measures the number of substitutions per site that have occurred along the evolutionary process on that edge.

Consider a transition matrix M_e ruling the evolutionary process on an edge $e : u \to v$ of a tree T and we assume the distribution at u is uniform. Then, Barry and Hartigan (1987) proved that, if M_e is not far from the identity matrix and under some mild assumptions, one can approximate the branch length of e by

$$l(e) = -\frac{1}{4} log|det(M_e)|. \tag{1.12}$$

Therefore the determinant of a matrix is an inverse measure for the amount of elapsed substitutions between the original and the final sequences. Note that as M_e is a stochastic matrix, then $|det(M_e)| \in [0, 1]$ (see Theorem 1.2.15). Then a matrix M_e with $|det(M_e)|$ close to zero represents a high amount of substitutions on e and therefore a large branch length on e. Conversely, a matrix M_e with determinant close to one means a short branch length on e.

When we are given two DNA sequences s_1, s_2, both descending from a common ancestor, it is more difficult to estimate the amount of substitutions that have occurred through the whole evolutionary process. One way to estimate this is via the *paralinear distance* (Lake, 1994).

Definition 1.1.33. Consider two DNA sequences s_1 and s_2. Let F be the 4×4 matrix whose (i, j) entry is the frequency of sites at which the observation at sequence s_0 is i and is j at sequence s_1. Let F_i denote the diagonal matrices whose diagonals are the frequency

vectors of nucleotides in s_i. Then the *paralinear distance* between s_1 and s_2 is defined by

$$d_{par}(s_1, s_2) = -log \frac{|det(F)|}{\sqrt{det(F_1)det(F_2)}} \tag{1.13}$$

and is assumed to be infinity if $det(F) = 0$.

Lake proved that under certain mild assumptions, the paralinear distance estimates the number of substitutions per site that occurred if s_1, s_2 had descended from a common ancestor. See also Section 3.4 for a generalization of this result. Moreover this measure is additive in the sense that, if two evolutionary processes are concatenated $u \to v \to w$, then $d_{par}(u, w) = d_{par}(u, v) + d_{par}(v, w)$.

The paralinear distance is not a mathematical distance since it does not satisfy the triangle inequality and $d(s_1, s_2) = 0$ does not imply $s_1 = s_2$. However the expression (1.13) defines a dissimilarity map:

Definition 1.1.34. (Semple and Steel, 2003). For any set X, a function $\delta : X \times X \to \mathbb{R}$ is called *dissimilarity map* if for all $x, y \in X$ $\delta(x, x) = 0$ and $\delta(x, y) = \delta(y, x)$.

Consider a dissimilarity map δ, then we say that δ is a *tree metric* if if there exists a tree T such that for all pairs of leaves x, y, $\delta(x, y)$ equals the length of the unique path from x to y. The four-point condition gives a characterization of a tree metric:

Theorem 1.1.35 (Four-point condition)**.** *A dissimilarity map δ is a tree metric if and only if it satisfies the triangle inequality and for any four leaves x, y, z, t, the maximum of the three numbers $\delta(x, y) + \delta(z, t)$, $\delta(x, z) + \delta(y, t)$ and $\delta(x, t) + \delta(y, z)$ is attained at least twice.*

For a proof see Theorem 2.36 in Pachter and Sturmfels (2005), for example.

1.2 Background on Algebra and Geometry

1.2.1 Matrices

Matrices are a basic object that play an important role in this work and shall be appearing throughout the following chapters. Here we introduce some notations, definitions and results that will be needed later.

Notation 1.2.1. We denote by $\mathcal{M}_n(\mathbb{K})$ the set of square $n \times n$ matrices with entries in a field $\mathbb{K}$ and by $\mathcal{M}_{n\times m}(\mathbb{K})$ the set of $n \times m$ matrices with entries in $\mathbb{K}$. Given a matrix

$M \in \mathcal{M}_n(\mathbb{K})$, $M(i,j)$ denotes the entry of M in the i-th row and j-th column. Usually $\mathbb{K}$ is equal to $\mathbb{R}$ or $\mathbb{C}$.

Notation 1.2.2. The vectors $\{e_1, \dots, e_n\}$ denote the standard basis of $\mathbb{K}^n$. Given a vector $v \in \mathbb{K}^n$ we denote by $v_i \in \mathbb{K}$ its components with respect to the standard basis, that is $v = \sum_i v_i e_i$. We usually identify a vector v with its coordinates in the standard basis $(v_1, \dots, v_n)^t$. Given a vector $v \in \mathbb{K}^n$, $diag(v)$ denotes the $n \times n$ matrix whose off diagonal entries are zero and whose (i,i) entry equals the i-th component of v.

Definition 1.2.3. Given a real square matrix $M \in \mathcal{M}_n(\mathbb{R})$, the *spectrum* of M is the set of all its eigenvalues and is denoted by $\Lambda(M)$. The *inertia* (or *signature*) of a real symmetric matrix S is the number of positive, negative and zero eigenvalues counted with multiplicity. It can be denoted by the triplet (i_+, i_-, i_0) where i_+ and i_- are also known as the *positive* and *negative inertia indices* of S, respectively.

Symmetric matrices have important properties regarding their spectrum. By the spectral theorem (see page 517 in Meyer, 2000), the eigenvalues of any real symmetric matrix S are all real and S diagonalizes through an orthonormal basis of eigenvectors.

Definition 1.2.4. A symmetric matrix $S \in \mathcal{M}_n(\mathbb{R})$ is said to be *positive semi-definite* (*PSD* for short) if $v^t S v$ is non-negative for any $v \in \mathbb{R}^n$; S is *positive definite* if for any $0 \neq v \in \mathbb{R}^n$ the product $v^t S v$ is strictly positive.

The following theorem allows us to characterize the positive semi-definite and positive definite matrices in terms of their eigenvalues.

Theorem 1.2.5 (Sylvester's law of inertia)**.** *Let $M, N \in \mathcal{M}_n(\mathbb{R})$ be two symmetric matrices. Then, M and N have the same inertia if and only if there exists an invertible matrix $H \in \mathcal{M}_n(\mathbb{R})$ such that $N = H^t M H$. As a consequence, a symmetric matrix $S \in \mathcal{M}_n(\mathbb{R})$ is positive semi-definite (respectively positive definite) if and only if all its eigenvalues are non-negative (respectively, positive).*

Moreover, positive definite and positive semi-definite matrices can also be characterized in terms of their entries as follows.

Theorem 1.2.6 (Sylvester's criterion)**.** *Let $S \in \mathcal{M}_n(\mathbb{R})$ be a symmetric matrix. Then, S is positive definite if and only if all its leading principal minors are positive and it is PSD if and only if every principal minor is non-negative.*

Theorem 1.2.7 (Jacobi, 1857)**.** *Let Δ_i be the i-th leading principal minor of a symmetric matrix $S \in \mathcal{M}_n(\mathbb{R})$ and consider the sequence $\Delta_0 = 1, \Delta_1, \dots, \Delta_n = det(M)$. Then, the number i_+ of positive eigenvalues of S is equal to the number of pairs of leading principal minors (Δ_{k-1}, Δ_k) $(k = 1, \dots, n)$ with the same sign.*

A proof for the *Sylvester's law of inertia* and the *Sylvester's criterion* can be found in the book by Horn and Johnson (1985) and the reader is refereed to Ghys and Ranicki (2016) for a proof of Theorem 1.2.7.

A common problem in applied linear algebra is finding the nearest matrix X to a given matrix M subject to some specific restrictions on X. For example, in many applications of different areas such as machine learning and statistics, the matrix M is obtained from empirical or simulated data and it has to be approximated by another matrix which is known to lie in some statistical model. Among the usual constraints required for X we may find: having a certain rank, being orthogonal, symmetric, positive-definite... We refer the reader to the paper of Higham (1989) for a nice introduction to this kind of applied problems. The quoted paper also contains a survey on theoretical results and computational methods usually applied to nearness problems for fundamental matrix properties like symmetry, positive-definiteness, orthogonality or normality.

In order to compute distances between matrices it is necessary to have a matrix norm defined; in this work we shall use the Frobenius norm which is equivalent to the euclidean norm of a matrix if understood as a vector. The following two definitions can be found in Horn and Johnson (1985).

Definition 1.2.8. Given a matrix $M \in \mathcal{M}_n(\mathbb{R})$ we define its *Frobenius norm* as

$$\|M\|_F = \sqrt{\sum_{i,j=1}^{n} M(i,j)^2}.$$

We begin by characterizing the distance of a matrix to a set of low rank matrices. To do so, we need to recall the definition of the singular values of a matrix.

Definition 1.2.9. The *singular values* of a matrix $M \in \mathcal{M}_{m\times n}(\mathbb{R})$ are denoted by $\sigma_i(M)$ and are the square roots of the positive eigenvalues of M^tM. If $\lambda_1 \geq \cdots \geq \lambda_r > 0$ are the positive eigenvalues of M^tM (with repetition if necessary), then

$$\sigma_i(M) = \sqrt{\lambda_i}.$$

The number r of singular values of a matrix coincides with its rank. If $\sigma_1(M) \geq \ldots \geq \sigma_r(M) > 0$ are the singular values of $M \in \mathcal{M}_{m\times n}(\mathbb{R})$, the *singular value decomposition* (SVD) of M is a factorization $M = UDV^t$ where $U \in \mathcal{M}_m(\mathbb{R})$, $V \in \mathcal{M}_n(\mathbb{R})$ are orthogonal matrices and $D \in \mathcal{M}_{m\times n}(\mathbb{R})$ has entries $D(i,i) = \sigma_i(M)$ if $i \leq r$, and 0 everywhere else.

The Frobenius norm of Definition 1.2.8 can also be expressed in terms of the singular values of M, $\|M\|_F = \sqrt{\sum_{i=1}^{n} \sigma_i(M)^2}$ as the Frobenius norm is invariant by the action of orthogonal matrices.

Theorem 1.2.10 (Eckart and Young, 1936)**.** *Let $\mathcal{M}_{\leq k}$ be the determinantal variety of $n \times n$ matrices with rank less than or equal to k. Then, given a matrix $M \in \mathcal{M}_n(\mathbb{R})$ with singular values $\sigma_1(M) \geq \cdots \geq \sigma_r(M)$, the distance from M to $\mathcal{M}_{\leq k}$ is*

$$\delta_k(M) := \min_{N \in \mathcal{M}_{\leq k}} \|M - N\|_F = \sqrt{\sum_{i=k+1}^{r} \sigma_i(M)^2}.$$

The problem of finding the closest symmetric matrix S to a real matrix $M \in \mathcal{M}_n(\mathbb{R})$ was solved by Fan and Hoffman (1955) and the solution is based on the decomposition of $\mathcal{M}_n(\mathbb{R})$ into the direct sum of the orthogonal subspaces of symmetric and skew-symmetric matrices.

Theorem 1.2.11 (Fan and Hoffman, 1955)**.** *Let $M \in \mathcal{M}_n(\mathbb{R})$ be a real square matrix, let $S = \dfrac{M + M^T}{2}$ be the "symmetric part" of M and $C = \dfrac{M - M^T}{2}$ be the "skew-symmetric part", so that $M = S + C$. Then, S is the closest symmetric matrix to M and the distance from M to the set of symmetric matrices is*

$$\delta_{sym}(M) := \|M - S\|_F = \|C\|_F.$$

Moreover, Higham (1988) characterized which is the closest positive semidefinite matrix to a given $M \in \mathcal{M}_n(\mathbb{R})$. The following result uses the *polar decomposition* UH of a matrix in which U is a orthogonal matrix and H is a positive semi-definite matrix.

Theorem 1.2.12 (Lemma 2.4 on Higham, 1988)**.** *Let $M \in \mathcal{M}_n(\mathbb{R})$ be a real square matrix, let S be the symmetric part of M and C be the skew-symmetric part as in Theorem 1.2.11. If $S = UH$ is the polar decomposition of S, then*

$$psd(M) := \frac{S + H}{2}$$

is the nearest positive semidefinite matrix to M (in the Frobenius norm). Moreover, if $\lambda_i(S)$ are the eigenvalues of S, then the Frobenius distance from $psd(M)$ to M is given by

$$\delta_{psd}(M) := \|M - psd(M)\|_F = \sqrt{\sum_{\lambda_i(S)<0} \lambda_i(S)^2 + \|C\|_F^2}.$$

We finish this section with some notation and results regarding stochastic matrices.

Notation 1.2.13. We write $\mathcal{H}_n$ for the hyperplane

$$\mathcal{H}_n = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid x_1 + \dots + x_n = 1\} \subset \mathbb{R}^n \tag{1.14}$$

and $\Delta^n := \left\{(x_1, \dots, x_n) \in \mathbb{R}^n \mid \sum_i x_i = 1, x_i \geq 0\right\}$ for the standard simplex in $\mathbb{R}^n$. Note that Δ^n is contained in $\mathcal{H}_n$.

Definition 1.2.14. We denote by $\mathbb{M}_n$ the set of $n \times n$ real matrices with rows summing to one,

$$\mathbb{M}_n = \left\{ M \in \mathcal{M}_n(\mathbb{R}) \mid \sum_j M(i,j) = 1 \ \forall i \right\},$$

and by $\mathbb{M}_n^+$ the set of stochastic matrices

$$\mathbb{M}_n^+ = \{M \in \mathbb{M}_n \mid M(i,j) \geq 0 \ \forall i, j\}.$$

Theorem 1.2.15 (Perron, 1907; Frobenius, 1912)**.** *Any 4×4 transition matrix M has an eigenvalue $\lambda_1 = 1$ with associated eigenvector $u = (1,1,1,1)^t$, and the other eigenvalues satisfy $|\lambda_i| \leq 1$. Moreover, if M is positive, then 1 is the only eigenvalue λ with $|\lambda| = 1$, it has algebraic multiplicity one and there exists a unique distribution $\overline{\pi}$ such that $\overline{\pi}^t M = \overline{\pi}^t$.*

1.2.2 Tensors

In this section we give a brief introduction to tensors and describe basic concepts and operations that shall be needed later. The main motivation for introducing tensors and other concepts of multilinear algebra is that the joint distribution at the leaves of a phylogenetic tree can be naturally seen as a tensor. This algebraic point of view is really useful and plays a key role in this work.

Consider n vector spaces $\mathbb{K}^{k_1}, \dots, \mathbb{K}^{k_n}$ with bases $\mathcal{B}_1, \dots, \mathcal{B}_n$. Then, we consider the tensor product $\mathbb{K}^{k_1} \otimes \cdots \otimes \mathbb{K}^{k_n}$, which is a vector space with natural basis $x_1 \otimes \cdots \otimes x_n$ with $x_i \in \mathcal{B}_i$. A $k_1 \times \cdots \times k_n$-tensor p is an element of $\mathbb{K}^{k_1} \otimes \cdots \otimes \mathbb{K}^{k_n}$. We will denote its entries in the natural basis as $p_{x_1,\dots,x_n} \in \mathbb{K}$. If $k := k_1 = \cdots = k_n$ we write $\mathcal{W} := \mathbb{K}^k$, and $\mathcal{W}^{\otimes n}$ for the tensor product $\mathcal{W} \otimes \overset{n)}{\dots} \otimes \mathcal{W}$. If $\mathcal{B} := \{v_1, \dots, v_k\}$ is a basis of $\mathcal{W}$, then the natural basis of $\mathcal{W}^{\otimes n}$ is $\{v_1 \otimes \cdots \otimes v_1, v_1 \otimes \cdots \otimes v_2, \dots, v_k \otimes \cdots \otimes v_k\}$. An n-tensor is an element

$$p = \sum_{x_1 \ldots, x_n \in \mathcal{B}} p_{x_1,\ldots,x_n}\, x_1 \otimes \ldots \otimes x_n \in \mathcal{W}^{\otimes n}.$$

We keep this notation for the rest of this section.

Definition 1.2.16. Consider two matrices $M_1 \in \mathcal{M}_{a_1 \times b_1}(\mathbb{K})$ and $M_2 \in \mathcal{M}_{a_2 \times b_2}(\mathbb{K})$. Then, the *tensor* or *Kronecker product* $M_1 \otimes M_2$ is a matrix in $\mathcal{M}_{a_1 a_2 \times b_1 b_2}(\mathbb{K})$ whose rows are indexed by pairs (i_1, i_2) and columns by (j_1, j_2) with $i_k \in [a_k]$ and $j_l \in [b_l]$. These pairs are considered in the lexicographic order and the entries of the resulting matrix $M_1 \otimes M_2$ are defined as follows: the entry $((i_1, i_2), (j_1, j_2))$ of $M_1 \otimes M_2$ corresponds to the product of the (i_1, j_1) entry of M_1 and the (i_2, j_2) entry of M_2:

$$(M \otimes N)((i_1, i_2), (j_1, j_2)) = M_1(i_1, j_1)\, M_2(i_2, j_2).$$

We proceed to define an action of matrices on tensors.

Definition 1.2.17. We write $\mathbb{M}_k^*$ for the subset of non-singular matrices whose rows sum to one, $\mathbb{M}_k^* = \{M \in \mathbb{M}_k \mid \det(M) \neq 0\}$. It can be seen that $\mathbb{M}_k^*$ with the matrix multiplication is a group. Following Sumner et al. (2017) we call it the *Markov group*.

Note that $\times^n \mathbb{M}_k^*$ is also a group with the operation induced by the operation of $\mathbb{M}_k^*$ componentwise. Based on Sumner et al. (2008), the *Markov action* is the group action of $\times^n \mathbb{M}_k^*$ on the space of tensors $\mathcal{W}^{\otimes n}$ defined by

$$\begin{array}{rccc} \mathcal{A}: & \mathcal{W}^{\otimes n} \times (\times^n \mathbb{M}_k^*) & \longrightarrow & \mathcal{W}^{\otimes n} \\ & (q, M_1, \ldots, M_n) & \longmapsto & q \cdot M_1 \otimes \cdots \otimes M_n, \end{array} \tag{1.15}$$

where the tensor $p := q \cdot M_1 \otimes \cdots \otimes M_n$ is defined in coordinates by

$$p_{x_1,\ldots,x_n} = \sum_{y_1,\ldots,y_n \in \mathcal{B}} q_{y_1,\ldots,y_n} M_1(y_1, x_1) \cdots M_n(y_n, x_n),$$

and the entries of M_i are indexed by the elements of the basis. The operation $q \cdot M_1 \otimes \cdots \otimes M_n$ coincides with the usual product of q seen as a row vector with the matrix $M_1 \otimes \cdots \otimes M_n$, i.e.,

$$p = \left(q_{v_1 \cdots v_1}, \cdots, q_{v_k \cdots v_k}\right)\, M_1 \otimes \cdots \otimes M_n.$$

Now, we define the product of a tensor by a vector or a matrix.

Definition 1.2.18. Given an n-tensor $p \in \mathcal{W}^{\otimes n}$, an integer $i \in \{1, \dots n\}$ and a vector $u \in \mathcal{W}$, we define a $(n-1)$-tensor $p *_i u$ as follows,

$$(p *_i u)_{x_1,\dots,x_{i-1},x_{i+1},\dots,x_n} = \sum_{x_i \in \mathcal{B}} p_{x_1,\dots,x_i,\dots,x_n} u_{x_i}.$$

In particular, if u equals $\mathbf{1} := (1, \dots, 1)^t$, we define the *$i$-th marginalization of p* as $p_{\dots+\dots} := p *_i \mathbf{1}$. Given a $k \times k$ matrix M, we define the *n-tensor $p *_i M$* as

$$p *_i M := p \cdot (Id \otimes \cdots \otimes Id \otimes \underset{\widehat{i}}{M} \otimes Id \otimes \cdots \otimes Id). \tag{1.16}$$

The components of $p *_i M$ can be expressed as

$$(p *_i M)_{x_1,\dots,x_n} = \sum_{y \in \mathcal{B}} p_{x_1,\dots,x_{i-1},y,x_{i+1},\dots,x_n} M(y, x_i). \tag{1.17}$$

Remark 1.2.19. Observe that the tensor matrix multiplication of (1.16) has the following two properties. If $M,\ N \in \mathbb{M}_k^*$ and $i, j \in [n]$ then

1. $(p *_i M) *_j N = \big(p *_j N\big) *_i M$, if $i \neq j$,
2. $(p *_i M) *_i N = p *_i (MN)$.

Definition 1.2.20. (Sumner et al., 2008) A *Markov invariant* $f(p)$ is a function on the entries of p such that for any p and any $g = M_1 \otimes \cdots \otimes M_n \in \times^n \mathbb{M}_k^*$

$$f(p \cdot g) = f(p \cdot M_1 \otimes \cdots \otimes M_n) = \lambda_g + f(p),$$

where $\lambda_g \in \mathbb{C}$ satisfies the additive group homomorphism property $\lambda_g + \lambda_{g'} = \lambda_{gg'}$ for all $g, g' \in \times^n \mathbb{M}_k^*$. Equivalently, if we want the multiplicative group homomorphism property on λ_g, we require $f(p \cdot g) = \lambda_g \cdot f(p)$ and $\lambda_g \lambda_{g'} = \lambda_{gg'}$.

We illustrate these definitions with an example.

Example 1.2.21. Let p be a 3-tensor in $\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$. We denote its entries by $p_{x,y,z}$ with $x, y, z \in \{1, 2\}$, where $p_{x,y,z}$ is the coefficient of $e_x \otimes e_y \otimes e_z$, and $\mathcal{B} = \{e_1, e_2\}$ is the standard basis of $\mathbb{C}^2$. If $u = (u_1, u_2) \in \mathbb{C}^2$, then the entries of the 2-dimensional tensor $p *_2 u$ are

$$(p *_2 u)_{1,1} = p_{1,1,1}u_1 + p_{1,2,1}u_2, \qquad (p *_2 u)_{1,2} = p_{1,1,2}u_1 + p_{1,2,2}u_2,$$
$$(p *_2 u)_{2,1} = p_{2,1,1}u_1 + p_{2,2,1}u_2, \qquad (p *_2 u)_{2,2} = p_{2,1,2}u_1 + p_{2,2,2}u_2.$$

The *first marginalization* $p_{+..} = p *_1 \mathbf{1}$ has components $(p_{+..})_{x,y} = p_{1,x,y} + p_{2,x,y}$.

Finally, given a matrix $M \in \mathcal{M}_2(\mathbb{C})$, the components of $q = p *_2 M$ are

$$\begin{aligned}
q_{1,1,1} &= p_{1,1,1} \cdot M(1,1) + p_{1,2,1} \cdot M(2,1), & q_{1,1,2} &= p_{1,1,2} \cdot M(1,1) + p_{1,2,2} \cdot M(2,1),\\
q_{1,2,1} &= p_{1,1,1} \cdot M(1,2) + p_{1,2,1} \cdot M(2,2), & q_{2,1,1} &= p_{2,1,1} \cdot M(1,1) + p_{2,2,1} \cdot M(2,1),\\
q_{2,2,1} &= p_{2,1,1} \cdot M(1,2) + p_{2,2,1} \cdot M(2,2), & q_{2,1,2} &= p_{2,1,2} \cdot M(1,1) + p_{2,2,2} \cdot M(2,1),\\
q_{1,2,2} &= p_{1,1,2} \cdot M(1,2) + p_{1,2,2} \cdot M(2,2), & q_{2,2,2} &= p_{2,1,2} \cdot M(1,2) + p_{2,2,2} \cdot M(2,2).
\end{aligned}$$

The polynomial $\sum_{x,y,z\in\{1,2\}} p_{x,y,z}$ is a Markov invariant with $\lambda_g = 1$ (for the multiplicative version) for all $g = M_1 \otimes M_2 \otimes M_3 \in \times^3 \mathbb{M}_2^*$, because all the rows of M_i sum to one and $\sum_{x,y,z\in\{1,2\}} (p \cdot g)_{x,y,z} = \sum_{x,y,z\in\{1,2\}} p_{x,y,z}$.

Remark 1.2.22. Note that we can consider 2-tensors as $k \times k$ matrices via the isomorphism

$$\begin{array}{rcl} \mathcal{W} \otimes \mathcal{W} & \longleftrightarrow & \mathcal{M}_k(\mathbb{K}) \\ p = \sum_{x,y\in\mathcal{B}} p_{x,y}\, x \otimes y & \longleftrightarrow & P = (p_{x,y})_{x,y\in\mathcal{B}}, \end{array}$$

where the rows of the matrix are indexed by the first component and columns by the second.

Moreover, any tensor p in $\mathcal{W}^{\otimes n}$ can be thought as a 2-tensor as follows. Consider the tensor product $\mathcal{W}^{\otimes n}$ and in order to make each vector space apparent denote $\mathcal{W}^{\otimes n} = \mathcal{W}_1 \otimes \mathcal{W}_2 \otimes \cdots \otimes \mathcal{W}_n$ (where each $\mathcal{W}_i = \mathcal{W}$). Let $A|B$ be a split of $[n]$, that is $A, B \subset [n]$ are ordered non-empty sets such that $A \cup B = [n]$ and $A \cap B = \emptyset$. Denote by a and b the respective cardinals of A and B and write $A = \{i_1, \dots, i_a\}$ and $B = \{j_1, \dots, j_b\}$. Then we define the vector spaces $\mathcal{W}_A := \mathcal{W}_{i_1} \otimes \cdots \otimes \mathcal{W}_{i_a}$ and $\mathcal{W}_B := \mathcal{W}_{j_1} \otimes \cdots \otimes \mathcal{W}_{j_b}$ with respective natural bases $\mathcal{B}_A$ and $\mathcal{B}_B$ obtained from the basis $\mathcal{B}$. Then there is a natural isomorphism between $\mathcal{W}_1 \otimes \dots \otimes \mathcal{W}_n$ and $\mathcal{W}_A \otimes \mathcal{W}_B$ given by:

$$\begin{array}{rcl} \mathcal{W}_1 \otimes \dots \otimes \mathcal{W}_n & \longleftrightarrow & \mathcal{W}_A \otimes \mathcal{W}_B \\ p = \sum_{x_i\in\mathcal{B}} p_{x_1,\dots,x_n}\, x_1 \otimes \cdots \otimes x_n & \longleftrightarrow & p = \sum_{\substack{x\in\mathcal{B}_A \\ y\in\mathcal{B}_B}} \tilde{p}_{x,y}\, x \otimes y, \end{array}$$

where $\tilde{p}_{x,y}$ is the entry $p_{x_1,\dots,x_n}$ of p for the corresponding vectors $x_i \in \mathcal{B}$. We illustrate this in Example 1.2.24.

Definition 1.2.23. Consider a split $A|B$ of $[n]$. The *flattening of p with respect to $A|B$* is the matrix $Flat_{A|B}(p)$ defined via the isomorphism

$$\begin{array}{rccc} Flat_{A|B}: & \mathcal{W}_1 \otimes \cdots \mathcal{W}_n \cong \mathcal{W}_A \otimes \mathcal{W}_B & \longleftrightarrow & \mathcal{M}_{k^a \times k^b}(\mathbb{K}) \\ & p = \sum_{\substack{x \in \mathcal{B}_A \\ y \in \mathcal{B}_B}} p_{x,y}\, x \otimes y & \longleftrightarrow & P = \left(p_{x,y}\right)_{x,y}, \end{array}$$

where the rows of P are ordered according to the elements of $\mathcal{B}_A$ and the columns according to the elements of $\mathcal{B}_B$.

Example 1.2.24. Consider a tensor $p \in \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$ as in Example 1.2.21. Denote it as $p = \sum_{x,y,z \in \{1,2\}} p_{x,y,z}\, e_x \otimes e_y \otimes e_z$ where $\mathcal{B} = \{e_1, e_2\}$ is the standard basis of $\mathbb{C}^2$.

Consider the split $A = \{1\}$ and $B = \{2,3\}$. Then $\mathcal{W}_A = \mathbb{C}^2$, $\mathcal{W}_B = \mathbb{C}^2 \otimes \mathbb{C}^2$ and their corresponding bases are $\mathcal{B}_A = \mathcal{B}$ and $\mathcal{B}_B = \{v_1 = e_1 \otimes e_1, v_2 = e_1 \otimes e_2, v_3 = e_2 \otimes e_1, v_4 = e_2 \otimes e_2\}$. Then, p can be seen as

$$\begin{aligned} p = & \sum_{x,y,z \in \{1,2\}} p_{x,y,z} e_x \otimes e_y \otimes e_z = \\ & = p_{1,1,1}\, e_1 \otimes (e_1 \otimes e_1) + p_{1,1,2}\, e_1 \otimes (e_1 \otimes e_2) + p_{1,2,1}\, e_1 \otimes (e_2 \otimes e_1) + p_{1,2,2}\, e_1 \otimes (e_2 \otimes e_2) + \\ & + p_{2,1,1}\, e_2 \otimes (e_1 \otimes e_1) + p_{2,1,2}\, e_2 \otimes (e_1 \otimes e_2) + p_{2,2,1}\, e_2 \otimes (e_2 \otimes e_1) + p_{2,2,2}\, e_2 \otimes (e_2 \otimes e_2), \end{aligned}$$

or equivalently as

$$\begin{aligned} p = \sum_{x \in [2],\, y \in [4]} \tilde{p}_{x,y} e_x \otimes v_y = & \tilde{p}_{1,1}\, e_1 \otimes v_1 + \tilde{p}_{1,2}\, e_1 \otimes v_2 + \tilde{p}_{1,3}\, e_1 \otimes v_3 + \tilde{p}_{1,4}\, e_1 \otimes v_4 + \\ & + \tilde{p}_{2,1}\, e_2 \otimes v_1 + \tilde{p}_{2,2}\, e_2 \otimes v_2 + \tilde{p}_{2,3}\, e_2 \otimes v_3 + \tilde{p}_{2,4}\, e_2 \otimes v_4, \end{aligned}$$

where $\tilde{p}_{i,1} = p_{i,1,1}$, $\tilde{p}_{i,2} = p_{i,1,2}$, $\tilde{p}_{i,3} = p_{i,2,1}$ and $\tilde{p}_{i,4} = p_{i,2,2}$ for $i \in \{1,2\}$.

The corresponding flattening is

$$Flat_{1|23}(p) = \begin{pmatrix} p_{1,1,1} & p_{1,1,2} & p_{1,2,1} & p_{1,2,2} \\ p_{2,1,1} & p_{2,1,2} & p_{2,2,1} & p_{2,2,2} \end{pmatrix}.$$

Lemma 1.2.25. *Consider a tensor q in $\mathcal{W}^{\otimes n}$ and n matrices $M_i \in \mathbb{M}_k^*$. Let $A = \{i_1, \ldots, i_a\}$ and $B = \{j_1, \ldots, j_b\}$ be a split of $[n]$. Then,*

$$Flat_{A|B}(q \cdot (M_1 \otimes \cdots \otimes M_n)) = \left(M_{i_1} \otimes \cdots \otimes M_{i_a}\right)^t Flat_{A|B}(q) \left(M_{j_1} \otimes \cdots \otimes M_{j_b}\right).$$

Proof. By generalizing the definition of the Kronecker product of two matrices we have

$$M_{i_1}(y_{i_1}, x_{i_1})M_{i_2}(y_{i_2}, x_{i_2}) \cdots M_{i_a}(y_{i_a}, x_{i_a}) = (M_{i_1} \otimes \cdots \otimes M_{i_a})(\mathbf{y}_a, \mathbf{x}_a).$$

Let $p = q \cdot (M_1 \otimes \cdots \otimes M_n)$ and denote $M_A := M_{i_1} \otimes \cdots \otimes M_{i_a}$ and $M_B = M_{j_1} \otimes \cdots \otimes M_{j_b}$.

Consider the entry $p_{x_1,\ldots,x_n}$ of p mapped to the entry $((x_{i_1}, \ldots, x_{i_a}), (x_{j_1}, \ldots, x_{i_b}))$ of $Flat_{A|B}(p)$ (where $\{x_{i_1}, \ldots, x_{i_a}, x_{j_1}, \ldots, x_{j_b}\} = \{x_1, \ldots, x_n\}$). For simplicity, denote $\mathbf{x}_a := (x_{i_1}, \ldots, x_{i_a})$ and $\mathbf{x}_b := (x_{j_1}, \ldots, x_{j_b})$ (and analogously for $\mathbf{y}_a$ and $\mathbf{y}_b$). Then

$$\begin{aligned}
&Flat_{A|B}(p)(\mathbf{x}_a, \mathbf{x}_b) = p_{x_1,\ldots,x_n} = \sum_{y_1,\ldots,y_n \in \mathcal{B}} q_{y_1,\ldots,y_n} M_1(y_1, x_1) \cdots M_n(y_n, x_n) = \\
&= \sum_{\substack{\mathbf{y}_a \in \mathcal{B}_A \\ \mathbf{y}_b \in \mathcal{B}_B}} Flat_{A|B}(q)(\mathbf{y}_a, \mathbf{y}_b) \left(M_{i_1}(y_{i_1}, x_{i_1}) \cdots M_{i_a}(y_{i_a}, x_{i_a})\right) \left(M_{j_1}(y_{j_1}, x_{j_1}) \cdots M_{j_b}(y_{j_b}, x_{j_b})\right) \\
&= \sum_{\mathbf{y}_a \in \mathcal{B}_A, \mathbf{y}_b \in \mathcal{B}_B} Flat_{A|B}(q)(\mathbf{y}_a, \mathbf{y}_b)\, M_A(\mathbf{y}_a, \mathbf{x}_a)\, M_B(\mathbf{y}_b, \mathbf{x}_b) = \\
&= \sum_{\mathbf{y}_a \in \mathcal{B}_A, \mathbf{y}_b \in \mathcal{B}_B} M_A^t(\mathbf{x}_a, \mathbf{y}_a)\, Flat_{A|B}(q)(\mathbf{y}_a, \mathbf{y}_b)\, M_B(\mathbf{y}_b, \mathbf{x}_b),
\end{aligned}$$

which corresponds to the $(\mathbf{x}_a, \mathbf{x}_b)$ entry of $M_A^t Flat_{A|B}(q) M_B$. □

1.2.3 Basic notions on Algebraic Geometry

In this section we recall some concepts and results regarding algebraic geometry that are needed in the sequel. Our introduction to this field follows the book by Cox, Little, and O'Shea (2007) and the reader is referred to it for further details on the topic. We shall work over a field $\mathbb{K}$ that later will be taken as either $\mathbb{R}$ or the complex field $\mathbb{C}$.

Affine algebraic varieties

Definition 1.2.26. An affine *algebraic variety* $\mathcal{V}$ in $\mathbb{K}^n$ is the common zero set of a collection of polynomials $f_1, \ldots, f_r \in \mathbb{K}[x_1, \ldots, x_n]$, this is, $\mathcal{V} = \mathcal{V}(f_1, \ldots, f_r)$ where

$$\mathcal{V}(f_1, \ldots, f_r) = \{p \in \mathbb{K}^n \mid f_1(p) = \cdots = f_r(p) = 0\}. \tag{1.18}$$

In this case we say that the set $\{f_1, \ldots, f_r\}$ *defines* $\mathcal{V}$ *set-theoretically*. The algebraic varieties in $\mathbb{K}^n$ are the closed sets of a topology known as the *Zariski topology*.

Instead of considering a set of polynomials we can consider the ideal I generated by this set of polynomials, $I = \langle f_1, \ldots, f_r \rangle$. Then, it is straightforward to see that $p \in \mathcal{V}(f_1, \ldots, f_r)$ if and only if $f(p) = 0\ \forall f \in I$. Thus we can denote $\mathcal{V}(I) = \mathcal{V}(f_1, \ldots, f_r)$.

Note that the definition of $\mathcal{V}(I)$ does not depend on the particular polynomials that we choose to generate I.

It is well known that every ideal $I \subseteq \mathbb{K}[x_1, ..., x_n]$ can be generated by a finite set of polynomials $f_1, \dots, f_m$ (this is the Hilbert's basis theorem); moreover one can start from a set in $\mathbb{K}^n$ and consider the polynomials that vanish on it:

Lemma 1.2.27. *Given any subset S of $\mathbb{K}^n$, the set of polynomials vanishing at every point $p \in S$ forms an ideal in $\mathbb{K}[x_1, \dots x_n]$ called the* ideal of S *and denoted by $I(S)$. Moreover, $\mathcal{V}(I(S))$ is the smallest algebraic variety containing S.*

Definition 1.2.28. The *Zariski closure* $\bar{S}$ of a subset S of $\mathbb{K}^n$ is the smallest algebraic variety containing S, and it is equal to $\mathcal{V}(I(S))$.

Note that if $\mathcal{V} = \mathcal{V}(f_1, \dots, f_r)$, then $\langle f_1, \dots, f_r \rangle \subseteq I(\mathcal{V})$ but in general the equality does not hold. Hilbert's Nullstellensatz states that if $\mathbb{K} = \mathbb{C}$, then $I(\mathcal{V})$ is the radical ideal of $\langle f_1, \dots, f_r \rangle$ (see Theorem 2 in 4.1, Cox, Little, and O'Shea, 2007).

Elimination Theory and Implicitization

Below we introduce basic concepts and results on *elimination theory* that shall be necessary in Chapter 2. We follow Chapter 3 of the book written by of Cox, Little, and O'Shea (2007) to give an introduction to this topic. The main goal of elimination theory is the study of algorithmic methods to eliminate some variables from systems of polynomial equations in order to solve them. The foundations of elimination theory are based on two main results, the *elimination theorem* and the *extension theorem*. We only present here the elimination theorem.

Definition 1.2.29. Given an ideal $I = \langle f_1, \dots, f_r \rangle \subset \mathbb{K}[x_1, \dots, x_n]$, the *elimination ideal* I_k is the ideal of $\mathbb{K}[x_{k+1}, \dots, x_n]$ defined by $I_k := I \cap \mathbb{K}[x_{k+1}, \dots, x_n]$.

Eliminating variables $x_1, \dots, x_k$ means finding a system of generators of I_k, that is, finding polynomials in I only with variables $x_{k+1}, \dots, x_n$ that generate I_k. Then, the first step is to have a procedure to obtain the elements of I_k. This follows from the following result and from the existence of Gröbner bases of an ideal (for a definition of Gröbner basis see Definition 5 on Chapter 2, Section 5 in Cox, Little, and O'Shea, 2007).

Theorem 1.2.30 (Elimination theorem)**.** *Let I be an ideal and let G be a Gröbner basis of I with respect to the lexicographical order with $x_1 > \cdots > x_n$. Then $G_k := G \cap \mathbb{K}[x_{k+1}, \dots, x_n]$ is a Gröbner basis of the k-th elimination ideal I_k for each k.*

Consider the variety $\mathcal{V}(I_k) \subset \mathbb{K}^{n-k}$ of an elimination ideal I_k. A point $\tilde{p} = (p_{k+1}, \dots, p_n) \in \mathcal{V}(I_k)$ is called a *partial solution* to $\mathcal{V}(I)$ and the goal of the extension step is to decide whether it can be extended to a solution $p \in \mathcal{V}(I)$. We shall deal with this problem in Chapter 2 by verifying ad hoc which solutions of the varieties $\mathcal{V}(I_k)$ can be extended to $\mathcal{V}(I)$.

If we start with a variety $\mathcal{V} = \mathcal{V}(I) \subseteq \mathbb{C}^n$ then we can obtain partial solutions by projecting it into $\mathbb{C}^{n-k}$:

$$\begin{array}{rccc} \pi_k : & \mathbb{C}^n & \longrightarrow & \mathbb{C}^{n-k} \\ & p = (p_1, \dots, p_n) & \longmapsto & (p_{k+1}, \dots, p_n). \end{array} \tag{1.19}$$

The following result guarantees that for any point $p \in \mathcal{V}$, the projection $\pi_k(p)$ lies in $\mathcal{V}(I_k)$.

Theorem 1.2.31 (Closure theorem)**.** *Write $\mathcal{V} = \mathcal{V}(I)$. Then $\mathcal{V}(I_k)$ is the Zariski closure of $\pi_k(\mathcal{V})$. Moreover, if $\mathcal{V} \neq \emptyset$, there exists an affine algebraic variety $\mathcal{W} \subset \mathcal{V}(I_k)$ such that $\mathcal{V}(I_k) \setminus \mathcal{W} \subseteq \pi_k(\mathcal{V}) \subseteq \mathcal{V}(I_k)$.*

The second part of the theorem says that even if $\pi_k(\mathcal{V})$ may not fill the whole $\mathcal{V}(I_k)$, the missing part lies in a strictly smaller variety.

Consider a set of polynomials $g_1, \dots, g_n \in \mathbb{K}[t_1, \dots, t_m]$. A map $\phi : \mathbb{K}^m \to \mathbb{K}^n$ is a *polynomial map* if all its components are polynomial functions. That is,

$$\begin{array}{rcl} \phi : \mathbb{K}^m & \longrightarrow & \mathbb{K}^n \\ (t_1, \dots, t_m) & \longmapsto & (g_1(t_1, \dots, t_m), \dots, g_n(t_1, \dots, t_m)), \end{array}$$

is a polynomial map and we say that $\phi(\mathbb{K}^m) \subset \mathbb{K}^n$ is a subset of $\mathbb{K}^n$ *parametrized* by ϕ. If $\mathcal{V}$ is an algebraic variety parametrized by ϕ, that is, the Zariski closure of Im ϕ, then a natural question is how to find polynomials $f_1, \dots, f_r$ defining $\mathcal{V}$. This is known as the *Implicitization Problem*. We do not deal explicitly with this problem in this thesis, although this has been deeply studied in the context of phylogenetics as we may see in Section 1.3.1.

Elimination theory can be of some help to solve the implicitization problem as follows. Consider the variety

$$\overline{\mathcal{V}} = \mathcal{V}\big(x_1 - g_1(t_1, \dots, t_m), \dots, x_n - g_n(t_1, \dots, t_m)\big) \subseteq \mathbb{K}^{n+m}$$

which is formed by the points $(t_1, \dots, t_m, g_1(t_1, \dots, t_m), \dots, g_n(t_1, \dots, t_m)) \in \mathbb{K}^{n+m}$. Given $t = (t_1, \dots, t_m)$, consider the two maps

$$\begin{array}{rcl} i : \mathbb{K}^m & \longrightarrow & \mathbb{K}^{n+m} \\ t & \longmapsto & (t, g_1(t), \dots, g_n(t)) \end{array} \quad \text{and} \quad \begin{array}{rcl} \pi_m : \mathbb{K}^{n+m} & \longrightarrow & \mathbb{K}^n \\ (t, x_1 \dots, x_n) & \longmapsto & (x_1 \dots, x_n). \end{array}$$

As $\phi = \pi_m \circ i$, it can be shown that $\phi(\mathbb{K}^m) = \pi_m(i(\mathbb{K}^m)) = \pi_m(\overline{\mathcal{V}})$, and so elimination theory can be used to determine the smallest variety containing $\mathcal{V} = \phi(\mathbb{K}^m)$:

Theorem 1.2.32. *Follow the same notation as before and denote by I the ideal $I := \langle x_1 - g_1(t), \dots, x_n - g_n(t)\rangle \subset \mathbb{K}[t_1, \dots, t_m, x_1 \dots, x_n]$ and I_m the elimination ideal $I_m = I \cap \mathbb{K}[x_1 \dots, x_n]$. Then $\mathcal{V}(I_m)$ is the smallest algebraic variety in $\mathbb{K}^n$ containing $\phi(\mathbb{K}^m)$.*

Euclidean distance degree

Another problem we shall encounter in Chapter 2 is, given a data point $p \in \mathbb{R}^n$, finding the closest point $\hat{p} \in \mathcal{V}$ with respect to the Euclidean distance. In what follows we give a brief introduction to this topic but we refer to the work by Draisma et al. (2016) for further details from a computational algebraic geometry perspective.

Consider an irreducible algebraic variety $\mathcal{V}$ with ideal $I(\mathcal{V}) = \langle f_1, \dots, f_r\rangle$. Then, the dimension of $\mathcal{V}$, $dim\,\mathcal{V}$, can be defined of as the maximum $d \in \mathbb{N}$ such that there exist a chain $V_0 \subsetneq \mathcal{V}_1 \subsetneq \cdots \subsetneq \mathcal{V}_d = \mathcal{V}$ of irreducible varieties.

Definition 1.2.33. Let $d_p(x)$ be the *squared Euclidean distance* from a given point $p \in \mathbb{R}^n$ to a point $x \in \mathbb{R}^n$,

$$d_p(x) = \sum_{i=1}^{n} (p_i - x_i)^2,$$

and define the *squared Euclidean distance from a point p to an algebraic variety $\mathcal{V} \subseteq \mathbb{R}^n$* as

$$d(p, \mathcal{V}) = \min_{x \in \mathcal{V}} d_p(x). \tag{1.20}$$

In order to find $d(p, \mathcal{V})$, we need to find a minimum of d_p (viewed as a function of x) restricted to $\mathcal{V}$. To this end, we extend the function d_p to the complex field, $d_p : \mathbb{C}^n \to \mathbb{C}$ and consider the critical points of d_p restricted to $\mathcal{V}$. In order to avoid computational problems, we want to exclude the singular points of $\mathcal{V}$ and consider them afterwards if necessary. A point of $\mathcal{V}$ that minimizes the distance to p will be one of these critical points. Recall that $p \in \mathcal{V}$ is singular if rank $J_{\mathcal{V}}(p) < c$ where $J_{\mathcal{V}}(p)$ is the Jacobian matrix formed by the partial derivatives $\partial_{x_j}(f_i)$ evaluated at p and c is the

codimension of $\mathcal{V}$ ($c = n - dim\,\mathcal{V}$). Denote by $\mathcal{V}_{sing}$ the singular locus of $\mathcal{V}$, i.e. the set of singular points of $\mathcal{V}$.

Lemma 1.2.34. *For a generic $p \in \mathbb{C}^n$, the number of critical points of d_p on $\mathcal{V} \setminus \mathcal{V}_{sing}$ is finite and constant on a dense open subset of $\mathbb{R}^n$.*

This quantity is called the *Euclidean distance degree* of $\mathcal{V}$ (*EDdegree* for short). We refer to Draisma et al. (2016) for an algorithm that computes the EDdegree. This algorithm can be adapted when $\mathcal{V}$ is defined by a polynomial map, as indicated in the quoted paper. In Chapter 2 we use `Magma` to compute the EDdegree of certain varieties. Once the EDdegree is computed, given a point $p \in \mathbb{R}^n$ (and assuming that is general enough) the distance $d(p, \mathcal{V})$ can be found by using numerical software that computes all critical points of $d_p(x)$ (we use the EDdegree to ensure that we obtain all critical points), selecting those with real coordinates, and evaluating the distance function to choose the minimum.

1.3 Algebraic Phylogenetics

1.3.1 Phylogenetic algebraic varieties

Lake (1987) and Cavender and Felsenstein (1987) introduced, in two independent works, the notion of *phylogenetic invariants*, polynomials that arise from a phylogenetic tree and a particular model. Moreover, they envisioned the use of these polynomials in phylogenetic reconstruction. In this section, we give an introduction to this approach.

Let T be a rooted phylogenetic tree and $\mathcal{M}$ be a nucleotide substitution model. Let r be the root of T and consider a Markov process on T as in Section 1.1.2, where the substitution parameters satisfy the constraints of the model $\mathcal{M}$.

As we have seen in Section 1.1.2, the components of the joint distribution p at the leaves are polynomials in the model parameters, so we can associate to each tree T a polynomial map $\phi_T^{\mathcal{M}} : \mathbb{R}^d \to \mathbb{R}^{4^n}$ that maps any d-tuple of parameters to a distribution vector of the 4^n possible observations at the leaves as follows:

$$\begin{array}{rcl} \phi_T^{\mathcal{M}} : \mathbb{R}^d & \longrightarrow & \mathbb{R}^{4^n} \\ \left(\pi; \{M_e\}_{e\in E(T)}\right) & \longmapsto & p = (p_{\mathtt{A},\mathtt{A}\ldots,\mathtt{A}}, p_{\mathtt{A},\mathtt{A}\ldots,\mathtt{C}}, p_{\mathtt{A},\mathtt{A}\ldots,\mathtt{G}}, \ldots, p_{\mathtt{T},\mathtt{T}\ldots,\mathtt{T}}), \end{array} \tag{1.21}$$

where d is the number of free substitution parameters according to the model assumed (see Section 1.1.3), and each coordinate $p_{x_1,\ldots,x_n}$ is expressed in terms of the distribution at the root π and the transition matrices M_e according to the expression (1.3).

Remark 1.3.1. Notice that, in order to interpret the parameters as probabilities, we need to restrict to non-negative real numbers. Analogously, the points in the image of $\phi_T^{\mathcal{M}}$ represent a distribution only if they lie in the standard Δ^{4^n} simplex. However, in order to use techniques from algebraic geometry, we abandon temporarily these constraints and work over the real field with no restrictions. In this case, although $p \in \operatorname{Im}\phi_T^{\mathcal{M}}$ may not be a distribution, its components always sum to one, i.e. $\operatorname{Im}\phi_T^{\mathcal{M}} \subset \mathcal{H}_{4^n}$. Indeed,

$$\sum_{x_1,\dots,x_n} p_{x_1,\dots,x_n} = \sum_{x_1,\dots,x_n} \sum_{x_r\in\Sigma} \sum_{\substack{x_v\in\Sigma \\ v\in \mathrm{Int}(T)}} \pi_{x_r} \prod_{e\in E(T)} M_e\left(x_{a(e)}, x_{d(e)}\right) =$$

$$= \sum_{x_r\in\Sigma} \sum_{x_v\in\Sigma} \pi_{x_r} \prod_{e\in E(T)} M_e\left(x_{a(e)}, x_{d(e)}\right) = \sum_{x_r\in\Sigma} \pi_{x_r} = 1, \tag{1.22}$$

where the first equality in (1.22) follows since the rows of each M_e sum up to 1.

Definition 1.3.2. We consider real parameters in general, and we refer to a set $\left\{\pi, \{M_e\}_{e\in E(T)}\right\}$ as *non-singular* parameters if

(i) At every node u of T the distribution of the random variable X_u is strictly positive.

(ii) The matrix M_e of every edge e is non-singular, that is, $M_e \in \mathbb{M}_4^*$.

Moreover, a set of parameters $\left\{\pi, \{M_e\}_{e\in E(T)}\right\}$ is *stochastic* if all the components of π and the entries of the transition matrices M_i are real numbers between 0 and 1. For stochastic parameters condition (i) above is equivalent to requiring that the distribution at the root π_r has no zero entry if (ii) holds.

Notation 1.3.3. For a quartet T, we shall denote $p := \phi_T(\pi; M_1, M_2, M_3, M_4, M_5)$ where M_5 is the matrix at the interior edge of the tree and each M_i with $i \neq 5$ is the transition matrix attached at the exterior edge adjacent to leaf i.

It is known that the image of the parametrization map $\phi_T^{\mathcal{M}}$ does not depend on the root position:

Theorem 1.3.4. *Let T be a phylogenetic tree rooted at r. Suppose we root it at a different node r' and call this tree T'. Then, for any set of parameters $\pi, \{M_e\}_{e\in E(T)}$ on a model $\mathcal{M}$ such that $\pi > 0$ and no column of any M_e is zero, there exist parameters $\pi', \{M'_e\}_{e\in E_{T'}}$ on the same model such that*

$$\phi_T^{\mathcal{M}}\left(\pi; \{M_e\}_e\right) = \phi_{T'}^{\mathcal{M}}\left(\pi'; \{M'_e\}_e\right).$$

This means that the root position within the tree cannot be inferred from the joint distribution at the leaves. This phenomenon is usually known as the *non-identifiability* of

the root position. Allman and Rhodes (2003b) proved this result for the general Markov model, however, the same proof can be adapted to group-based or other type of models (see Casanellas, Fernández-Sánchez, and Michałek (2017)).

Example 1.3.5. Consider the quartet $T = T_{12|34}$ with the root r placed at the parent node of leaves 1 and 2 and with distribution $\pi > 0$ at r. Assume M_5 is the stochastic GM matrix at the interior edge and M_i with $i = 1, \dots, 4$ are the stochastic GM matrices attached to the edges adjacent to the leaves. Consider now the quartet T' where the root r' is located at the parent node of leaves 3 and 4. If r' follows the distribution $\pi' = M_5^t \pi$ with $\pi' > 0$, and the stochastic GM matrix $M_5' = diag(\pi')^{-1} M_5^t diag(\pi)$ is attached at the interior edge of T', then

$$\phi_T(\pi; M_1, M_2, M_3, M_4, M_5) = \phi_{T'}(\pi'; M_1, M_2, M_3, M_4, M_5').$$

This equality can be deduced from Allman and Rhodes (2004).

In order to state the following theorem we need to consider matrices *diagonal largest in column* (*DLC*, for short). A stochastic matrix $M \in \mathbb{M}_n^+$ is DLC if the largest entry in each column is placed at the diagonal, that is, $M(i,i) > M(j,i)$ for all $j \neq i$.

Theorem 1.3.6 (Chang, 1996)**.** *Let T be a phylogenetic tree with no nodes of degree 2. Suppose p is a distribution that has arisen on T with non-singular stochastic parameters $\pi, \{M_e\}_{e \in E(T)}$ and assume that each matrix M_i is DLC and different than the identity matrix. Then the full model is identifiable; in other words, from the joint distribution p, the tree topology of T, the distribution π and the transition matrices M_i are uniquely determined.*

If we forget the assumption of *DLC* and we just require transition matrices to be invertible and not a permutation matrix, then the topology of T is still identifiable and M_i are determined up to a label swapping in the sense of Allman and Rhodes (2004). For biological applications it will be convenient to assume that transition matrices are *DLC* in some cases since this leads to a more biologically realistic situation where the transition matrices should not be too far from the identity matrix.

As a consequence of Theorems 1.3.4 and 1.3.6, from now on, we deal with trees with no nodes of degree two. Even if we root the tree at an interior node in order to specify the substitution parameters, the position of this root does not play any role in practice.

Example 1.3.7. The first and second trees presented in Figure 1.4 are rooted phylogenetic trees, and the root is located at a node of degree two. If we unroot these trees (i.e., remove the root label and delete the degree 2 node by replacing the edges adjacent to it with a

single edge) then all three trees in the figure are isomorphic. In this thesis we shall be working with trivalent trees (as the one on the right).

Definition 1.3.8. The *phylogenetic variety* associated to a tree T and a model $\mathcal{M}$, denoted by $\mathcal{V}_T^{\mathcal{M}}$, is the smallest algebraic variety containing the image $\operatorname{Im}\phi_T^{\mathcal{M}}$ or equivalently, the Zariski closure of $\operatorname{Im}\phi_T^{\mathcal{M}}$. We will call $\phi_T^{\mathcal{M}}$ the parametrization map of $\mathcal{V}_T^{\mathcal{M}}$.

Definition 1.3.9. We define the *stochastic phylogenetic region* associated to a tree T and a model $\mathcal{M}$, denoted by $\left(\mathcal{V}_T^{\mathcal{M}}\right)^+$, as the image by $\phi_T^{\mathcal{M}}$ of the stochastic parameters:

$$\left(\mathcal{V}_T^{\mathcal{M}}\right)^+ = \left\{p \in \mathcal{V}_T^{\mathcal{M}} \mid p = \phi_T^{\mathcal{M}}\left(\pi; \{M_e\}_e\right), \text{ where } \pi \text{ and } M_i \text{ are stochastic}\right\}.$$

Remark 1.3.10. The ideal $I(\operatorname{Im}\phi_T^{\mathcal{M}})$ coincides with the ideal of the variety $\mathcal{V}_T^{\mathcal{M}}$ (see Lemma 1.2.27). We shall denote it by $I_T^{\mathcal{M}}$. As pointed out in Theorem 1.3.4, this variety is independent of the node chosen as root in T. As a consequence, two unrooted trees with the same tree topology (i.e. isomorphic phylogenetic trees) define the same phylogenetic variety.

Definition 1.3.11. Given a tree T with n leaves and an evolutionary model $\mathcal{M}$, the polynomials in $I_T^{\mathcal{M}}$ are called *phylogenetic invariants* of T. If f is a polynomial in $I_T^{\mathcal{M}}$ which does not belong to $I_{T'}^{\mathcal{M}}$ for some other tree $T' \in \mathcal{T}_n$, $T' \neq T$, then f is called a *topology invariant* of T.

We refer to the work by Jarvis, Holland, and Sumner (2013) for a good review on phylogenetic invariants. Finding phylogenetic invariants is not a trivial task; however, some of them are evident. By Remark 1.3.1, for any n-leaf tree and any model $\mathcal{M}$, $I_T^{\mathcal{M}}$ contains the trivial invariant:

$$\sum_{x_1,\ldots,x_n \in \Sigma} p_{x_1,\ldots,x_n} - 1.$$

The other invariants shall depend on T and $\mathcal{M}$. Below we give an example of some phylogenetic and topology invariants for a concrete tree and model.

Example 1.3.12. Let T be the quartet presented in Figure 1.6 and consider the Markov process under the Jukes Cantor model. From Example 1.1.26 we have

$$p_{\mathtt{A},\mathtt{A},\mathtt{A},\mathtt{A}} = p_{\mathtt{C},\mathtt{C},\mathtt{C},\mathtt{C}} = p_{\mathtt{G},\mathtt{G},\mathtt{G},\mathtt{G}} = p_{\mathtt{T},\mathtt{T},\mathtt{T},\mathtt{T}},$$
$$p_{\mathtt{A},\mathtt{A},\mathtt{A},\mathtt{C}} = p_{\mathtt{A},\mathtt{A},\mathtt{A},\mathtt{G}} = \cdots = p_{\mathtt{T},\mathtt{T},\mathtt{T},\mathtt{C}} = p_{\mathtt{T},\mathtt{T},\mathtt{T},\mathtt{G}}.$$

These equalities hold due to the fact that JC69 matrices are invariant under any permutation of the states $\{\mathtt{A}, \mathtt{C}, \mathtt{G}, \mathtt{T}\}$. Thus, these linear equations give rise to phylogenetic invariants ($p_{\mathtt{A},\mathtt{A},\mathtt{A},\mathtt{A}} - p_{\mathtt{C},\mathtt{C},\mathtt{C},\mathtt{C}}$, for example) but not topology invariants since they hold for any distribution arising on any $T \in \mathcal{T}_4$ with JC69 parameters. However, the following equalities found by Lake (1987):

$$p_{\mathtt{A},\mathtt{C},\mathtt{A},\mathtt{C}} + p_{\mathtt{A},\mathtt{C},\mathtt{G},\mathtt{T}} = p_{\mathtt{A},\mathtt{C},\mathtt{G},\mathtt{C}} + p_{\mathtt{A},\mathtt{C},\mathtt{A},\mathtt{T}}, \tag{1.23}$$

$$p_{\mathtt{A},\mathtt{C},\mathtt{C},\mathtt{A}} + p_{\mathtt{A},\mathtt{C},\mathtt{T},\mathtt{G}} = p_{\mathtt{A},\mathtt{C},\mathtt{C},\mathtt{G}} + p_{\mathtt{A},\mathtt{C},\mathtt{T},\mathtt{A}}, \tag{1.24}$$

give rise to topology invariants for $T_{12|34}$. Indeed, (1.23) does not hold for distributions that have arisen from $T_{13|24}$ and distributions on $T_{14|23}$ do not satisfy (1.24).

For some time, there was a certain belief that one could infer which tree best explains the data by directly evaluating phylogenetic invariants on the observed distributions and testing whether these evaluations were close to zero. In other words, suppose $\tilde{p}$ is an observed distribution, obtained by computing the observed relative frequencies of an alignment. Then, if $\tilde{p}$ has been obtained by N independent samples of a theoretical distribution p that has arisen on T under a model $\mathcal{M}$, and if N is large enough, $f(\tilde{p})$ should be close to zero for each invariant $f \in I_T^{\mathcal{M}}$. Therefore, the use of these invariants could produce a method to infer the tree topology without requiring to infer the numerical parameters of the model (see Lake, 1987; Casanellas and Fernández-Sánchez, 2007). However, it has been shown that the use of invariants for inferring the tree topologies has some difficulties and it has not a straightforward application to topology inference. Nevertheless, this approach is important to understand the motivation of this work.

Because of Hilbert's basis theorem, in order to obtain all phylogenetic invariants for a tree T and a model $\mathcal{M}$, it is enough to determine a finite set of generators for the ideal $I_T^{\mathcal{M}}$. Finding such set of generators is a complex problem and has been an active field of research in the last years. This can be seen as an implicitization problem in algebraic geometry: given a parametric representation of an algebraic variety $\mathcal{V}_T^{\mathcal{M}}$, with a polynomial parameterization map $\phi_T^{\mathcal{M}}$, find an implicit representation of it as the set of zeros of a collection of polynomials (see Section 1.2.3). This implicitization problem can be addressed computationally via the elimination procedure previously described. For simple models and small trees, this can be done using computational algebra software such as Macaulay2 (Grayson and Stillman, 2009) or Singular (Greuel, Pfister, and Schönemann, 2009). However, this is not feasible when the number of leaves on the tree or the number of parameters increases. Due to the complexity of the problem an alternative approach has been finding set-theoretic descriptions of the phylogenetic varieties (see for instance, Casanellas, Fernández-Sánchez, and Michałek, 2017; Draisma and Kuttler, 2009; Allman and

Rhodes, 2008). Nowadays this description is available for any tree and the most common models. However, to to follow this approach, it would be convenient that the polynomials defining the variety $\mathcal{V}_T^{\mathcal{M}}$ would have a biological interpretation. On the other hand, we do not need to describe the whole variety but only the stochastic phylogenetic region (see Casanellas, Fernández-Sánchez, and Michałek, 2017).

In Section 1.3.3 we show how to produce some phylogenetic invariants for the GM model called *edge invariants*, which will play a fundamental role in our work.

1.3.2 The tensor of joint distribution

In this section, we view the joint distribution at the leaves of an n-leaf phylogenetic tree (described in Section 1.1.2) as an n-tensor and explain how the Markov action acts on it.

Consider the vector space $\mathcal{W} := \mathbb{R}^4$ and identify the canonical basis of $\mathcal{W}$ with the set $\Sigma = \{\mathtt{A}, \mathtt{C}, \mathtt{G}, \mathtt{T}\}$. Then the natural basis of the tensor power $\mathcal{W}^{\otimes n}$ is $\{x_1 \otimes \ldots \otimes x_n\}_{x_1,\ldots,x_n \in \Sigma}$. For instance, the natural basis of $\mathcal{W} \otimes \mathcal{W} \otimes \mathcal{W}$ is $\{\mathtt{A} \otimes \mathtt{A} \otimes \mathtt{A},\ \mathtt{A} \otimes \mathtt{A} \otimes \mathtt{C}, \ldots,\ \mathtt{T} \otimes \mathtt{T} \otimes \mathtt{T}\}$.

The joint distribution $p = (p_{x_1,\ldots,x_n})_{x_1,\ldots,x_n \in \Sigma} \in \mathbb{R}^{4^n}$ can be thought as a n-tensor in $\mathcal{W}^{\otimes n}$ whose components in the natural basis above are $p_{x_1,\ldots,x_n}$, that is,

$$p = \sum_{x_1,\ldots,x_n \in \Sigma} p_{x_1,\ldots,x_n} x_1 \otimes \ldots \otimes x_n.$$

The following result illustrates the effect of applying a Markov action (see Definition 1.2.17) on such a tensor. Even though this can be done for any n-leaf phylogenetic tree, for simplicity we state the result for quartets. In the following lemma we see how, given a 4-tensor arising on a quartet T with certain substitution parameters, we obtain a new 4-tensor that arises on the same tree T but with different parameters.

Lemma 1.3.13. *Let $p = \phi_T(\pi; M_1, M_2, M_3, M_4, M_5)$ be a 4-tensor that has arisen on a quartet T. Consider the matrices N_1, N_2, N_3 and N_4 such that $N_i \in \mathbb{M}_4^*$ for each i. Then the resulting tensor after applying the Markov action*

$$p \longmapsto q := p \cdot (N_1 \otimes \ldots \otimes N_4)$$

is equal to $\phi_T(\pi; M_1N_1, M_2N_2, M_3N, M_4N_4, M_5)$.

Proof. We do the proof for $T = T_{12|34}$, as the other cases are analogous. By Definition 1.2.17 and (1.4) we have that

$$\begin{aligned}
q_{x_1,x_2,x_3,x_4} &= \sum_{y_i\in\Sigma} p_{y_1,y_2,y_3,y_4} N_1(y_1,x_1)N_2(y_2,x_2)N_3(y_3,x_3)N_4(y_4,x_4) \\
&= \sum_{y_i\in\Sigma} N_1(y_1,x_1)N_2(y_2,x_2)N_3(y_3,x_3)N_4(y_4,x_4)\cdot \\
&\quad \cdot \sum_{m,l\in\Sigma} \pi_m M_1(m,y_1)M_2(m,y_2)M_5(m,l)M_3(l,y_3)M_4(l,y_4) = \\
&= \sum_{m,l\in\Sigma}\sum_{y_i\in\Sigma} \pi_m \left(M_1(m,y_1)N_1(y_1,x_1)\right)\left(M_2(m,y_2)N_2(y_2,x_2)\right)\cdot \\
&\quad \cdot M_5(m,l)\left(M_3(l,y_3)N_3(y_3,x_3)\right)\left(M_4(l,y_4)N_4(y_4,x_4)\right) = \\
&= \sum_{m,l\in\Sigma} \pi_m \left(M_1N_1\right)(m,x_1)\left(M_2N_2\right)(m,x_2)M_5(m,l)\left(M_3N_3\right)(l,x_3)\left(M_4N_4\right)(l,x_4).
\end{aligned}$$

By Equation (1.4), q corresponds to the distribution that has arisen on $T_{12|34}$ with parameters $\{\pi, M_1N_1, M_2N_2, M_3N_3, M_4N_4, M_5\}$. □

In particular, we have:

Corollary 1.3.14. *Let $p = \phi_T(\pi; M_1, M_2, M_3, M_4, M_5)$ be a 4-tensor that has arisen on a quartet T. If M_i is non-singular for some i, then the tensor $q = p *_i (M_i^{-1}M)$ is the joint distribution arising on the same tree and with the same parameters as p except for M_i which has been replaced by M.*

Proof. Suppose $i = 1$, the computations for $i = 2, 3, 4$ are equivalent. By Definition 1.2.18, we have that $q = p *_1 (M_1^{-1}M) = (M_1^{-1}M \otimes Id \otimes Id \otimes Id)\cdot p$ and by Lemma 1.3.13, $q = \phi_T(\pi; M_1(M_1^{-1}M), M_2Id, M_3Id, M_4Id, M_5) = \phi_T(\pi; M, M_2, M_3, M_4, M_5)$. □

In the remaining of the section, we show some technical results related to marginalizations of tensors that arise from a general Markov process on a tree T. These results have been extracted from Allman, Rhodes, and Taylor (2014).

Lemma 1.3.15. *Let $p = \phi_{T_3}(\pi; M_1, M_2, M_3)$ be a 3-tensor that arises on the 3-leaf tree T_3 with matrices $M_i \in \mathbb{M}_4$ (see Figure 1.9 left). Then, the three possible marginalizations of p, $p_{..+} = p *_3 \mathbf{1}$, $p_{.+.} = p *_2 \mathbf{1}$ and $p_{+..} = p *_1 \mathbf{1}$ are*

$$\begin{aligned}
p_{..+} &= M_1^t diag(\pi) M_2, \\
p_{.+.} &= M_1^t diag(\pi) M_3, \\
p_{+..} &= M_2^t diag(\pi) M_3.
\end{aligned} \tag{1.25}$$

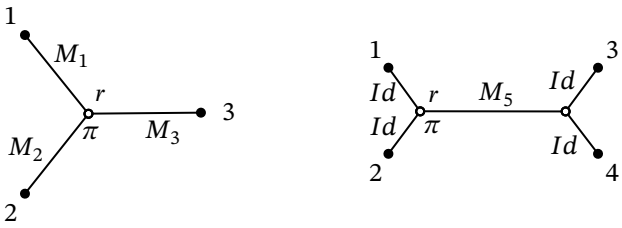


FIGURE 1.9: At the left: the 3-leaf trivalent tree $T_3 \in \mathcal{T}_3$ with substitution parameters $\{\pi; M_1, M_2, M_3\}$. At the right: the quartet $T_{12|34} \in \mathcal{T}_4$ with substitution parameters $\{\pi; Id, Id, Id, Id, M_5\}$.

Proof. We compute the 3rd marginalization of p, $p_{..+}$. From (1.3) we have that

$$(p_{..+})_{x_1,x_2} = (p *_3 \mathbf{1})(x_1,x_2) = \sum_{x_3\in\Sigma} 1 \cdot p_{x_1,x_2,x_3} = \sum_{x_3\in\Sigma}\sum_{i\in\Sigma} \pi_i M_1(i,x_1)M_2(i,x_2)M_3(i,x_3) =$$

$$= \sum_{i\in\Sigma} \pi_i M_1(i,x_1)M_2(i,x_2)\left(\sum_{x_3\in\Sigma} M_3(i,x_3)\right) = \sum_{i\in\sigma} \pi_i M_1(i,x_1)M_2(i,x_2).$$

The elements of this sum are written in terms of the vector π, the x_1-th column of M_1 and the x_2-th column of M_2. Equivalently, this is the product of the x_1-th row of M_1^t, the diagonal matrix $diag(\pi)$ and the x_2-th column of M_2. Therefore this expression for all x_1 and x_2 becomes

$$p_{..+} = p *_3 \mathbf{1} = M_1^t diag(\pi) M_2. \tag{1.26}$$

Similarly we can compute the expressions of $p_{..+}$ and $p_{+..}$. □

The above marginalizations extend naturally to tensors that come from quartets.

Lemma 1.3.16. *Let $p = \phi_T(\pi; M_1, M_2, M_3, M_4, M_5)$ be a tensor arising on the tree $T = T_{12|34}$ (see Figure 1.6). Then the marginalizations $p_{..+.}$ and $p_{...+}$ are:*

$$p_{..+.} = \phi_{T_3}(\pi; M_1, M_2, M_5M_4), \quad p_{...+} = \phi_{T_3}(\pi; M_1, M_2, M_5M_3), \tag{1.27}$$

and the double marginalizations $p_{+..+}$, $p_{+.+.}$, $p_{.+.+}$, and $p_{.++.}$ can be computed in terms of the transition matrices as follows:

$$\begin{aligned} p_{+..+} &= M_2^t diag(\pi) M_5 M_3, \quad p_{+.+.} = M_2^t diag(\pi) M_5 M_4, \\ p_{.+.+} &= M_1^t diag(\pi) M_5 M_3, \quad p_{.++.} = M_1^t diag(\pi) M_5 M_4. \end{aligned} \tag{1.28}$$

Proof. The proof of (1.27) is straightforward by summing over the corresponding index as in the previous lemma.

To compute the double marginalizations in (1.28), we need to marginalize the tensor over two different positions. We do the case $p_{+..+}$ since the other ones are analogous. Firstly we compute $p_{...+}$ and then we compute $(p_{...+})_{+..}$. From Lemma 1.3.16 we have that $p_{...+} = \phi_{T_3}(\pi; M_1, M_2, M_5M_3)$ and using Lemma 1.3.15, we obtain

$$p_{+..+} = (p_{...+})_{+..} = M_2^t diag(\pi) M_5 M_3.$$

□

Fourier coordinates

The discrete Fourier transform is a powerful tool to analyse properties of group-based models and the pattern distributions arising from them. This idea was introduced by Hendy (1989) and Hendy and Penny (1989) and was further explored by Székely, Steel, and Erdös (1993).

Let M be a K81 matrix and write $m_{\mathtt{A}}, m_{\mathtt{C}}, m_{\mathtt{G}}, m_{\mathtt{T}}$ and $\bar{\mathtt{A}}, \bar{\mathtt{C}}, \bar{\mathtt{G}}, \bar{\mathtt{T}}$ for the Fourier parameters of M and the Fourier basis, respectively (see Lemma 1.1.28).

Definition 1.3.17. For a tensor p in $(\mathcal{W})^{\otimes 4}$ we write $\bar{p} = (\bar{p}_{\mathtt{AAAA}}, \bar{p}_{\mathtt{AAAC}}, \dots, \bar{p}_{\mathtt{TTTG}}, \bar{p}_{\mathtt{TTTT}})^t$ for the coordinates of p in the basis $\left\{\bar{A} \otimes \bar{A} \otimes \bar{A} \otimes \bar{A}, ..., \bar{T} \otimes \bar{T} \otimes \bar{T} \otimes \bar{T}\right\}$ induced by the Fourier basis $\bar{\Sigma} = \left\{\bar{\mathtt{A}}, \bar{\mathtt{C}}, \bar{\mathtt{G}}, \bar{\mathtt{T}}\right\}$.

The relationship between the natural coordinates and the Fourier coordinates of p can be presented via the Markov action:

$$\bar{p} = p \cdot \left(H^{-1} \otimes H^{-1} \otimes H^{-1} \otimes H^{-1}\right) = \frac{1}{4^4} p \cdot (H \otimes H \otimes H \otimes H),$$

where H is the Hadamard matrix introduced in Lemma 1.1.28.

Remark 1.3.18. Since $\frac{1}{2}H$ is an orthogonal matrix, so is $U := \left(\frac{1}{2}H\right) \otimes \left(\frac{1}{2}H\right) \otimes \left(\frac{1}{2}H\right) \otimes \left(\frac{1}{2}H\right)$. Therefore, if $\|\cdot\|$ is the usual Euclidean norm, we have

$$\left\|\bar{p} - \bar{q}\right\|^2 = \left\|\frac{1}{2^4} p \cdot U - \frac{1}{2^4} q \cdot U\right\|^2 = \frac{1}{4^4} \|p - q\|^2,$$

and the Euclidean distance between tensors p and q can be computed in terms of their Fourier coordinates as $d(p, q) = 16||\bar{p} - \bar{q}||$.

If one considers the bijection between Σ and the group $G = (\mathbb{Z}_2 \times \mathbb{Z}_2, +)$ introduced in Remark 1.1.29, then the previous change of coordinates can be understood as the discrete Fourier transform on G^4. The following result states that the polynomial parametrization ϕ_T^{K81} becomes monomial in the Fourier parameters:

Theorem 1.3.19 (Evans and Speed, 1993). *Let p be a joint distribution arising on the tree $T = T_{A|B}$ with transitions matrices M_i in the K81 model. If $m^i_{\mathtt{A}}, m^i_{\mathtt{C}}, m^i_{\mathtt{G}}, m^i_{\mathtt{T}}$ are the Fourier parameters of M_i, then the Fourier coordinates of p are*

$$\overline{p}_{x_1x_2x_3x_4} = \begin{cases} \frac{1}{4^4} m^1_{x_1} m^2_{x_2} m^3_{x_3} m^4_{x_4} m^5_{\sum_{i\in A} x_i} & \text{if } \sum_{i\in A} x_i = \sum_{j\in B} x_j, \\ 0 & \text{otherwise,} \end{cases}$$

where the sum of elements in Σ is given by the bijection $\Sigma \leftrightarrow \mathbb{Z}_2 \times \mathbb{Z}_2$ of Remark 1.1.29.

Example 1.3.20. Consider the tree $T_{12|34}$ evolving under the JC69 model and let m_i be the eigenvalue of M_i different from one. Then for any three different states $x, y, z \neq \mathtt{A}$ the entries of $\overline{p}$ in Fourier coordinates satisfy

$$\begin{array}{llll} \overline{p}_{\mathtt{AAAA}} = 1 & \overline{p}_{\mathtt{A}x\mathtt{A}x} = m_2m_4m_5 & \overline{p}_{\mathtt{A}xyz} = m_2m_3m_4m_5 & \overline{p}_{xxxx} = m_1m_2m_3m_4 \\ \overline{p}_{\mathtt{AA}xx} = m_3m_4 & \overline{p}_{\mathtt{A}xx\mathtt{A}} = m_2m_3m_5 & \overline{p}_{x\mathtt{A}yz} = m_1m_3m_4m_5 & \overline{p}_{xxyy} = m_1m_2m_3m_4 \\ \overline{p}_{xx\mathtt{AA}} = m_1m_2 & \overline{p}_{x\mathtt{AA}x} = m_1m_4m_5 & \overline{p}_{xy\mathtt{A}z} = m_1m_2m_4m_5 & \overline{p}_{xyxy} = m_1m_2m_3m_4m_5 \\ & \overline{p}_{x\mathtt{A}x\mathtt{A}} = m_1m_3m_5 & \overline{p}_{xyz\mathtt{A}} = m_1m_2m_3m_5 & \overline{p}_{xyyx} = m_1m_2m_3m_4m_5 \end{array}$$

and any other entry equals zero. Therefore we can see $\overline{p}$ as a vector in $\mathbb{R}^{14}$ by considering these 14 coordinates. Moreover, observe that

$$\overline{p}_{xxxx} = \overline{p}_{xxyy} \qquad \text{and} \qquad \overline{p}_{xyxy} = \overline{p}_{xyyx}, \tag{1.29}$$

so $\overline{p}$ actually lies in the linear space $L_{T_{12|34}} \subset \mathbb{R}^{14}$ of dimension 12 defined by these two equations.

The equalities in (1.29) (which correspond to the Lake invariants of (1.23) and (1.24) in Fourier coordinates) are satisfied since $\overline{p}$ corresponds to the Fourier coordinates of a distribution p that has arisen on $T_{12|34}$. However, if we had taken $T = T_{13|24}$ then we would have $\overline{p}_{xxxx} = \overline{p}_{xyxy}$ and $\overline{p}_{xxyy} = \overline{p}_{xyyx}$. Then, the varieties L_T differ for each tree and their union spans a linear space of dimension 14 (Casanellas, Fernández-Sánchez, and Kedzierska, 2012).

1.3.3 Flattening matrix and edge invariants

In this section, given an split $A|B$ of $[n]$ we apply the flattening of Definition 1.2.23 to a distribution p arising on a tree. The flattening matrix will reveal important information for our purposes.

Let $A|B$ be a split of the leaves of a n-leaf tree T and let X_A and X_B be the joint random variables associated to A and B, namely $X_A := (X_x)_{x\in A}$ and $X_B := (X_y)_{y\in B}$. Then X_A and X_B can take $a := 4^{|A|}$ and $b := 4^{|B|}$ states respectively. Given a tensor $p \in \mathcal{W}^{\otimes n}$ the *flattening matrix* $Flat_{A|B}(p)$ is the $a \times b$ matrix whose entries are the joint distributions of the observations of X_A and X_B:

$$Flat_{A|B}(p) = \begin{array}{c} \\ \text{States of } X_A \end{array} \begin{array}{c} \text{States of } X_B \\ \begin{pmatrix} p_{x_1y_1} & p_{x_1y_2} & \cdots & p_{x_1y_b} \\ p_{x_2y_1} & p_{x_2y_2} & \cdots & p_{x_2y_b} \\ \vdots & \vdots & \ddots & \vdots \\ p_{x_ay_1} & p_{x_ay_2} & \cdots & p_{x_ay_b} \end{pmatrix} \end{array}.$$

Example 1.3.21. Consider a tensor p that has arisen on the quartet $T_{12|34}$ presented at Figure 1.6. Then, $Flat_{12|34}(p)$ is the 16×16 matrix:

$$Flat_{12|34}(p) = \begin{array}{c} \\ \text{States at leaves 1 and 2} \end{array} \begin{array}{c} \text{States at leaves 3 and 4} \\ \begin{pmatrix} p_{\texttt{AAAA}} & p_{\texttt{AAAC}} & p_{\texttt{AAAG}} & \cdots & p_{\texttt{AATT}} \\ p_{\texttt{ACAA}} & p_{\texttt{ACAC}} & p_{\texttt{ACAG}} & \cdots & p_{\texttt{ACTT}} \\ p_{\texttt{AGAA}} & p_{\texttt{AGAC}} & p_{\texttt{AGAG}} & \cdots & p_{\texttt{AGTT}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{\texttt{TTAA}} & p_{\texttt{TTAC}} & p_{\texttt{TTAG}} & \cdots & p_{\texttt{TTTT}} \end{pmatrix} \end{array}.$$

Theorem 1.3.22 (Allman and Rhodes, 2006)**.** *Let p be the tensor arising on the tree $T_{12|34}$ with transition matrix M_5 at the interior edge and matrices $M_i \in \mathbb{M}_4$ with $i \in [4]$ at the exterior edges, and let q be a tensor arising on the same tree with the same matrix M_5 at the interior edge and the identity matrix at the exterior edges, as shown in Figure 1.9 right. Then, for any split $ab|cd$ on $[4]$ the flattening matrices of p can be written as:*

$$Flat_{ab|cd}(p) = (M_a \otimes M_b)^t \, Flat_{ab|cd}(q) \, (M_c \otimes M_d). \tag{1.30}$$

Proof. By Lemma 1.3.13 we have that $p = q \cdot (M_1 \otimes M_2 \otimes M_3 \otimes M_4)$. Then, the result is a direct consequence of Lemma 1.2.25. □

Lemma 1.3.23. *With the same assumptions and notation of the previous theorem, the flattening matrices of q satisfy the next properties:*

(i) $Flat_{12|34}(q) = Flat_{12|43}(q) = Flat_{21|34}(q) = Flat_{21|42}(q)$.

(ii) $Flat_{13|24}(q)$ *is equal to* $Flat_{14|23}(q)$ *and coincides with the diagonal matrix whose diagonal entries are the* 16 *entries of* $diag(\pi)M_5$.

Proof. First of all, observe that the entries of q are

$$q_{x_1,x_2,x_3,x_4} = \begin{cases} \pi_{x_1} M_5(x_1,x_3) & \text{if } x_1 = x_2 \text{ and } x_3 = x_4, \\ 0 & \text{otherwise,} \end{cases} \tag{1.31}$$

Suppose $\{a,b\} = \{1,2\}$ and $\{c,d\} = \{3,4\}$. Then, the $((x_1,x_2),(x_3,x_4))$ entry of any $Flat_{ab|cd}(q)$ equals $\pi_{x_1} M_5(x_1,x_3)$ if $x_1 = x_2$ and $x_3 = x_4$, and is zero otherwise. This observation is sufficient to prove (i).

The equality of $Flat_{13|24}(q) = Flat_{14|23}(q)$ follows since $Flat_{13|24}(q)((x_1,x_2),(x_3,x_4)) = q_{x_1,x_3,x_2,x_4}$, $Flat_{14|23}(q)((x_1,x_2),(x_3,x_4)) = q_{x_1,x_3,x_4,x_2}$ and $q_{x_1,x_3,x_2,x_4} = q_{x_1,x_3,x_4,x_2}$ because of (1.31). Finally, the matrix $Flat_{13|24}(q)$ is diagonal since $Flat_{13|24}(q)((x_1,x_2),(x_3,x_4))$ is different from zero if and only if $(x_1,x_2) = (x_3,x_4)$, and by (1.31) the entries at the diagonal are $\pi_{x_1} M_5(x_1,x_3)$. □

Edge invariants

Theorem 1.3.24. *(Allman and Rhodes, 2006, 2008; Casanellas and Fernández-Sánchez, 2011) Let T be an n-leaf tree and consider $p = \phi_T\left(\pi, \{M_e\}_{e\in E(T)}\right)$. If $A|B$ is a non-trivial edge split of T, then $Flat_{A|B}(p)$ has rank ≤ 4 or, equivalently, the 5×5 minors of $Flat_{A|B}(p)$ vanish. Moreover if $A'|B'$ is an incompatible split of T then $Flat_{A'|B'}(p)$ has rank ≥ 16 generically. In particular, the 5×5 minors of $Flat_{A|B}(p)$ are topology invariants for T.*

Proof. We give the proof for quartets since this is the case we are interested in. We suppose $T = T_{12|34}$, as the proof for the other quartets is analogous. Suppose T is rooted at the left internal node r with distribution at the root π and with transition matrices M_i as shown in Figure 1.6. Let p be the joint distribution at the leaves of T. By Theorem 1.3.22, the flattening matrix $Flat_{12|34}(p)$ can be written as

$$Flat_{12|34}(p) = (M_1 \otimes M_2)^t \, Flat_{12|34}(q) \, (M_3 \otimes M_4),$$

where q is a distribution that has arisen on the tree $T_{12|34}$ with the identity matrix at the exterior edges, M_5 at the interior edge and with the distribution π at the root (see Figure 1.9 right).

We prove that $Flat_{12|34}(q)$ has rank at most 4 and consequently the rank of $Flat_{12|34}(p)$ is less than or equal to four. Recall that the components $q_{x_1x_2x_3x_4}$ of q are equal to zero if $x_1 \neq x_2$ or $x_3 \neq x_4$. Then, the flattening matrix $Flat_{12|34}(q)$ may have entries $Flat_{12|34}(q)((x_1,x_2),(x_3,x_4))$ different from zero only in the rows where $x_1 = x_2$ and columns where $x_3 = x_4$, that is

$$Flat_{12|34}(q) = \begin{pmatrix} p_{\mathtt{AAAA}} & 0\cdots 0 & p_{\mathtt{AACC}} & 0\cdots 0 & p_{\mathtt{AAGG}} & 0\cdots 0 & p_{\mathtt{AATT}} \\ 0 & & 0 & & 0 & & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & & 0 & & 0 & & 0 \\ p_{\mathtt{CCAA}} & 0\cdots 0 & p_{\mathtt{CCCC}} & 0\cdots 0 & p_{\mathtt{CCGG}} & 0\cdots 0 & p_{\mathtt{CCTT}} \\ \vdots & & \vdots & & \vdots & & \vdots \\ p_{\mathtt{GGAA}} & 0\cdots 0 & p_{\mathtt{GGCC}} & 0\cdots 0 & p_{\mathtt{GGGG}} & 0\cdots 0 & p_{\mathtt{GGTT}} \\ \vdots & & \vdots & & \vdots & & \vdots \\ p_{\mathtt{TTAA}} & 0\cdots 0 & p_{\mathtt{TTCC}} & 0\cdots 0 & p_{\mathtt{TTGG}} & 0\cdots 0 & p_{\mathtt{TTTT}} \end{pmatrix}. \quad (1.32)$$

Therefore, the entries different from zero of this matrix are concentrated in a 4×4 submatrix and so each 5×5 minor equals zero.

Consider now an incompatible split of T, $A=\{1,3\}$ and $B=\{2,4\}$. By Lemma 1.3.23, $Flat_{13|24}(q)$ is a diagonal matrix whose diagonal entries are those of $diag(\pi)M_5$. If M_5 and π are positive then $Flat_{13|24}(q)$ has rank 16. Moreover if M_1, M_2, M_3 and M_4 are of maximal rank then

$$Flat_{13|24}(p) = (M_1 \otimes M_3)^t \, Flat_{13|24}(q)\,(M_2 \otimes M_4)$$

has also rank 16. □

Remark 1.3.25. From the previous proof we see that if p is a tensor that has arisen on a quartet $T_{A|B}$ with positive and non-singular parameters, then the rank of $Flat_{A'|B'}(p)$ is 16 for any incompatible split $A'|B'$. Thus, in this case the word "generically" in the statement of Theorem 1.3.24 refers to positive and non-singular parameters.

Finally, observe that the invariants we found regarding the minors of these flattening matrices are associated with the split of taxa induced by the internal edge of the tree. For that reason they are known as *edge invariants*.

1.3.4 Stochastic Conditions

So far we have focused on the algebraic description of $\mathrm{Im}\phi_T^{\mathcal{M}}$. However, in addition to the polynomial equalities describing $\mathcal{V}_T^{\mathcal{M}}$, inequalities are expected to play a role in describing the joint distributions. Note that the space of probability distributions arising as the image of stochastic parameters on a tree can be notably smaller than the set of points on the variety. Then, inequalities are crucial in determining if a point in a phylogenetic variety lies in the stochastic region. In this section we recall some theoretical results that allow us to provide some conditions to ensure that joint distributions come from stochastic parameters.

We follow the work by Allman, Rhodes, and Taylor (2014) where the authors give necessary and sufficient conditions for tensors having arised on quartet trees with stochastic parameters. In this work we summarize them in the following result, which characterizes the stochastic region $\mathcal{V}_T^+$ for $T = T_{12|34}$.

Theorem 1.3.26 (Allman, Rhodes, and Taylor, 2014)**.** *Let $p = \phi_T(\pi; M_1, \dots, M_5)$ be a 4-tensor that has arisen on $T = T_{12|34}$ with non-singular parameters. Then, p arises from stochastic parameters if and only if the marginalizations $p_{+\dots}$ and $p_{\dots+}$ arise from stochastic parameters and the 16×16 matrix*

$$det(p_{+..+})det(p_{.+.+})Flat_{13|24}(p *_2 (adj(p^t_{+..+})p^t_{.+.+})) *_3 (adj(p_{.+.+})p_{.++.})) \tag{1.33}$$

is positive semidefinite.

The idea of the proof is as follows. Assume the root r of T is placed at the interior node near leaves 1 and 2 as it is shown in Figure 1.6. Let M_i, $i = 1, 2, 3, 4$ be associated to the edges leading to leaves, M_5 the matrix on the internal edge and π the distribution at the root. By assumption, the rows of these matrices sum to 1. We define the matrices

$$\begin{aligned} N_{32} &:= p^t_{+..+}, \quad N_{31} := p^t_{.+.+}, \\ N_{14} &:= p_{.++.}, \quad N_{13} := p_{.+.+}, \end{aligned}$$

which, by Lemma 1.3.16, correspond to the following matrix products:

$$\begin{array}{ll} N_{32} = M_3^t M_5^t diag(\pi) M_2, & N_{31} = M_3^t M_5^t diag(\pi) M_1, \\ N_{14} = M_1^t diag(\pi) M_5 M_2, & N_{13} = M_1^t diag(\pi) M_5 M_3. \end{array} \tag{1.34}$$

We define now the tensor $\tilde{p}$ that arises on the same parameters as p except that M_2 has been replaced by M_1 and M_4 has been replaced by M_3. By Lemma 1.3.13 and Corollary 1.3.14 this can be done by applying the following Markov action on p:

$$\tilde{p} = p \cdot \left(Id, N_{32}^{-1} N_{31}, N_{13}^{-1} N_{14}, Id\right) = p \cdot \left(Id, M_2^{-1} M_1, M_3^{-1} M_4 t, Id\right), \tag{1.35}$$

or, equivalently,

$$\tilde{p} = \left(p *_2 N_{32}^{-1} N_{31}\right) *_3 N_{13}^{-1} N_{14} = \left(p *_2 M_2^{-1} M_1\right) *_3 M_3^{-1} M_4.$$

Since $\tilde{p}$ arises from the same parameters than p except that M_2 has been replaced by M_1 and M_3 by M_4, we can write the 16×16 flattening matrix of the tensor $\tilde{p}$ as

$$Flat_{13|24}(\tilde{p}) = (M_1 \otimes M_4)^t D (M_1 \otimes M_4),$$

where D is the diagonal matrix that contains the 16 entries of $diag(\pi)M_5$ (see Theorem 1.3.22 and Lemma 1.3.23 (ii)).

Since by hypothesis the 3-marginalizations $p_{+\dots}$ and $p_{\dots+}$ arise from stochastic parameters, π is stochastic and M_1, M_2, M_3, and M_4, are stochastic matrices (see Theorem 3.6 in Allman, Rhodes, and Taylor, 2014). Moreover if p arises from non-singular parameters, M_1 and M_4 are non-singular and π has positive entries. Thus $M_1 \otimes M_4$ is also non-singular. Then as $Flat_{13|24}(p)$ is a symmetric matrix, it is positive semidefinite if and only if the diagonal entries of D are non-negative. Thus if $Flat_{13|24}(\tilde{p})$ is positive semidefinite, as π is positive, we can ensure that M_5 has non-negative entries. If we multiply $Flat_{13|24}(\tilde{p})$ by the square of the appropriate non-zero determinant we clear denominators and obtain the algebraic expression (1.33) stated in the theorem.

Then, the theoretical result that we have seen in this section, together with conditions on the marginalizations of p proved in Allman, Rhodes, and Taylor (2014), allow us to provide a complete description of the model, given by phylogenetic invariants and certain polynomial inequalities which characterize points that are the image of stochastic parameters.

That is, a distribution $p \in \mathbb{R}^{4^4}$ has arisen on $T_{12|34}$ with non-singular stochastic parameters if and only if

- the marginalizations $p_{+\dots}$ and $p_{\dots+}$ arise from stochastic parameters (or equivalently, satisfy conditions of Theorem 3.6 in Allman, Rhodes, and Taylor, 2014),
- the 5×5 minors of $Flat_{12|34}(p)$ are zero (or analogously rank $Flat_{12|34}(p) \leq 4$) and
- if $\tilde{p} = \left(p *_2 N_{32}^{-1}N_{31}\right) *_3 N_{13}^{-1}N_{14}$ then the matrix $Flat_{13|24}(\tilde{p})$ is positive semidefinite.

This result can be found in Theorem 4.6 of Allman, Rhodes, and Taylor (2014).

Recall that a subset of $\mathbb{R}^n$ is called *semi-algebraic set* if it is defined by a finite sequence of polynomial equations $P(x_1, \dots, x_n) = 0$ and inequalities $Q(x_1, \dots, x_n) > 0$, or it is a finite union of such sets. As a consequence of the Sylvester's criterion (see Theorem 1.2.6) which translates the conditions of matrices being positive definite/semidefinite in terms of semi-algebraic conditions on the entries of those matrices, we can view the stochastic phylogenetic region as a semi-algebraic set.

The tensor $\tilde{p}$ constructed in (1.35) arises from the same parameters that p except that M_2 has been replaced by M_1, and M_3 by M_4. Then $\tilde{p}$ is the joint distribution of the tree presented in Figure 1.10. Observe that this tree is symmetric with respect to the interior edge, so we can state the following result.

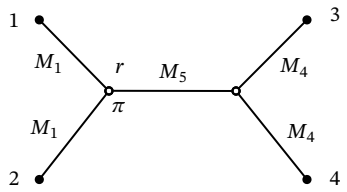


FIGURE 1.10: Quartet $T_{12|34}$ with substitution parameter $\{\pi; M_1, M_1, M_4, M_4, M_5\}$.

Corollary 1.3.27 (Garrote-López, 2016)**.** *Let p be a 4-tensor whose components sum to 1. Suppose that $p = \phi_T(\pi, M_1, M_2, M_3, M_4, M_5)$ with $T = T_{12|34}$. Let $\tilde{p}$ be constructed as in (1.35). Then,*

$$Flat_{13|24}(\tilde{p}) = Flat_{14|23}(\tilde{p}),$$

and

$$Flat_{12|34}(\tilde{p}) \neq Flat_{13|24}(\tilde{p}).$$

Proof. Using Theorem 1.3.22, Lemma 1.3.23 (ii), and the fact that $\tilde{p}$ is defined by replacing M_2 with M_1 and M_3 with M_4, we have

$$Flat_{13|24}(\tilde{p}) = (M_1 \otimes M_4)^t D(M_1 \otimes M_4) = Flat_{14|23}(\tilde{p}). \tag{1.36}$$

In contrast, if $\mathrm{q} = \phi_{12|34}(\pi; Id, Id, Id, Id, M_5)$, then

$$Flat_{12|34}(\tilde{p}) = (M_1 \otimes M_1)^t Flat_{12|34}(q)(M_4 \otimes M_4),$$

which is, in general, not equal to (1.36). □

1.4 PHYLOGENETIC RECONSTRUCTION METHODS

In the last decades, there has been a boom of new methods to infer phylogenetic trees from multiple sequence alignments. While some of them (*character-based methods*) directly use the columns of an alignment and the corresponding vector of observed relative frequencies, the so-called *distance-based methods* use pairwise information from the sequences in the alignment usually represented as a matrix of pairwise distances. On the other hand, whereas some methods are limited to reconstruct quartets (which can be used as input to build larger trees), others can deal with a larger number of sequence alignments.

As there is a unique trivalent tree with three leaves, the first interesting case is the reconstruction of quartets. In this case the input data is a sequence alignment on the set [4] of four taxa and the goal is to find which tree $T_{12|34}$, $T_{13|24}$ or $T_{14|23}$ fits the data best. To this end, phylogenetic reconstruction methods provide a score for each tree.

Sumner et al. (2017) suggested a criterion to decide when three scores $s_{ab|cd}$, $s_{ac|bd}$, $s_{ad|bc}$ obtained from data can be used to estimate the statistical confidence in the three quartet trees $T_{ab|cd}, T_{ac|bd}, T_{ad|bc}$. Suppose $F \in \Delta^{4^4}$ is the distribution obtained from N independent samples from the multinomial distribution p that has arisen on a tree T, we write $F \sim \mathrm{Mult}(N; p)$. If q is a frequency vector obtained from an alignment, we assume that it has been sampled from some distributions on a tree and N can be interpreted as the length of this alignment.

Definition 1.4.1. A *quartet inference measure* for F is a triplet

$$s(F) = \left(s_{12|34}(F), s_{13|24}(F), s_{14|23}(F)\right),$$

such that each $s_{ab|cd}(F)$ is a statistically interpretable confidence that F has arisen from N independent samples from some distribution p that has arisen on $T_{ab|cd}$. We say that $s(F)$ is a *normalized system of weights* if the scores are non-negative and sum to one.

Sumner et al. (2017) also presented a number of desirable statistical properties that any quartet inference measure should satisfy. In this work we consider the following two properties:

- *Property I:* If we permute the taxa $\{1,2,3,4\}$, then the scores $s_{12|34}(F)$, $s_{13|24}(F)$, $s_{14|23}(F)$ are permuted accordingly.
- *Property II (strong):* If $F \sim \text{Mult}(N;p)$, $g \in \mathbb{M}_4^*$ and $F' \sim \text{Mult}(N;p\cdot g)$ then, there exists a scalar λ_g such that $\mathbb{E}(s(F')) = \mathbb{E}(s(F)) + \lambda_g(1,1,1)$ where λ_g satisfies the additive group homomorphism property $\lambda_g + \lambda_{g'} = \lambda_{gg'}$ (depending on the definition of the quartet inference measure, this property can be equivalently stated in terms of a multiplicative group homomorphism property, see Definition 1.2.20).

Finally, we define the statistical consistency of a phylogenetic reconstruction method.

Definition 1.4.2. Consider a phylogenetic reconstruction method that given a distribution F outputs a phylogenetic tree. If $F \sim \text{Mult}(N;p)$ and p has arisen from a Markov process on the tree T, then we say that the method is *statistically consistent* if the probability that the method selects T as the correct tree tends to 1 as $N \to \infty$.

In what follows we briefly recall some phylogenetic reconstruction methods and for each one we present a weighting system for quartets. In Section 1.4.5 we present three quartet based methods that use these weights as input to reconstruct larger trees: Quartet Puzzling (QP), Weight Optimization (WO) and Willson's (WIL) method.

1.4.1 Maximum likelihood

Maximum likelihood is a statistical concept that has been widely used in many areas of biology. It was J. Felsenstein (1973) the pioneer on using it in phylogenetics. The basic idea of the *Maximum likelihood (`ML`)* methods is the following.

Given a vector F of observed relative frequencies obtained from an alignment of n taxa and a nucleotide substitution model $\mathcal{M}$, the likelihood function $\mathcal{L}_F(T;\theta)$ corresponds to the probability of observing such distribution F under the tree T with substitution parameters $\theta \in \mathcal{M}$. Then, `ML` attempts to estimate the tree topology $\hat{T} \in \mathcal{T}_n$ and the substitution parameters $\hat{\theta} \in \mathcal{M}$ that maximize the likelihood function $\mathcal{L}_F(T;\theta)$:

$$(\hat{T},\hat{\theta}) = \underset{\substack{T\in\mathcal{T}_n,\\ \theta\in\mathcal{M}}}{\arg\max}\{\mathcal{L}_F(T;\theta)\} = \underset{\substack{T\in\mathcal{T}_n,\\ \theta\in\mathcal{M}}}{\arg\max}\{Pr(F|T,\theta)\}.$$

The idea that relies on ML is simple but the optimization problem is computationally expensive, especially for large number of taxa taking into account that the number of phylogenetic trees grows more than exponentially in n (see Remark 1.1.10). Moreover, numerical optimization methods do not guarantee a global maximum nor the convergence of the method.

If we restrict to quartets so that $F \in \Delta^{4^4}$, we define a weighting system to ML according to Ranwez and Gascuel (2001): if $\mathcal{L}_F(T)$ is the likelihood for the tree T (that is $\mathcal{L}_F(T) = \max_{\theta\in\mathcal{M}} \mathcal{L}_F(T;\theta)$), then the normalized weights for the three quartets in $\mathcal{T}_4$ are defined by

$$\texttt{ML}(F) = \frac{1}{L}\left(\mathcal{L}_F(T_{12|34}), \mathcal{L}_F(T_{13|24}), \mathcal{L}_F(T_{14|23})\right).$$

where $L = \mathcal{L}_F(T_{12|34}) + \mathcal{L}_F(T_{13|24}) + \mathcal{L}_F(T_{14|23})$. ML is a statistically consistent method and if $F \sim \text{Mult}(N;\phi_T(\theta))$ with $T = T_{12|34}$, then $\texttt{ML}(F) \to (1,0,0)$ as $N \to \infty$.

1.4.2 Neighbor-joining algorithm

The Neighbor-joining algorithm (NJ) was proposed by Saitou and Nei (1987) and is the most used distance based method. Given n DNA sequences, the NJ algorithm takes as input the distances between any pair of taxa $i, j \in X$ given by a dissimilarity map $\delta : X \times X \to \mathbb{R}$ (see Definition 1.1.34) and outputs a phylogenetic tree.

One of the main advantages of the NJ algorithm (as well as other distance-based methods) is its computational speed (as it is of order n^3). Although the algorithm works for any dissimilarity map, we restrict to the case of the paralinear distance d_{par} of Definition 1.1.33.

Based on the work by Mihaescu, Levy, and Pachter (2009), we can define scores to weight quartets according to NJ. For the quartet $T_{ab|cd}$, consider

$$l_{ab|cd} = \frac{1}{4}(d_{par}(a,c) + d_{par}(a,d) + d_{par}(b,c) + d_{par}(b,d)) - \frac{1}{2}(d_{par}(a,b) + d_{par}(c,d)),$$

which represents the length of the interior edge of the tree $T = T_{ab|cd}$ if the distances were obtained from T. In order to assign a non-negative score for each topology, we define

$w_{ab|cd} = \exp\left(l_{ab|cd}\right)$. Then, the normalized weights for the three trees are taken as

$$\texttt{NJ}(F) = \frac{1}{w_{12|34} + w_{13|24} + w_{14|23}} \left(w_{12|34}, w_{13|24}, w_{14|23}\right).$$

If p has arisen on $T_{12|34}$ with some substitution parameters, then $w_{12|34}$ is the largest among the three weights (due to the four-point condition and the additivity of paralinear distances).

1.4.3 Maximum Parsimony and the Felsenstein zone

The Maximum Parsimony (MP for short) method for phylogenetic inference was introduced by Camin and Sokal (1965). This is a character-based method that computes the minimum number of nucleotide substitutions that are needed on a tree T to observe the given data. Then, MP outputs the tree with the least number of substitutions needed.

Felsenstein (1978) showed that maximum parsimony is not a statistically consistent method for some substitution parameters. The set of substitution parameters for which MP is inconsistent is known as the *Felsenstein zone*. The trees with parameters on this zone are subject to the *long branch attraction phenomenon* and consist on trees with two non-sister long branches and very short branch lengths (see Section 1.1.4) at the other edges, as illustrated in Figure 1.11.

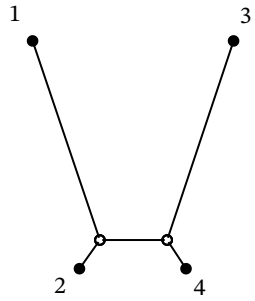


FIGURE 1.11: Tree subject to the long branch attraction phenomenon.

1.4.4 A rank-based method: Erik+2

Erik+2 (Fernández-Sánchez and Casanellas, 2016) is based on Theorem 1.3.24 that has been used in different phylogenetic reconstruction methods (see Allman, Kubatko, and Rhodes, 2016; Chifman and Kubatko, 2015, 2019).

Given a distribution $F \in \Delta^{4^4}$, one can compute the flattening matrix $Flat_{A|B}(F)$ and the distance of such matrix to the matrices of rank less than or equal to 4, $\delta_4(Flat_{A|B}(F))$

(see Theorem 1.2.10). Because of Theorem 1.3.24, if F has arisen from the tree $T_{A|B}$ this distance is zero. Instead of computing directly this distance, `Erik+2` computes for each quartet tree T the value

$$s_{A|B}(F) = \frac{\delta_4\left(Flat^r_{A|B}(F)\right) + \delta_4\left(Flat^c_{A|B}(F)\right)}{2}, \tag{1.37}$$

where $Flat^r_{A|B}(F)$ and $Flat^c_{A|B}(F)$ stand for the normalized matrices obtained by dividing the rows and columns (respectively) of the flattening matrix $Flat_{A|B}(F)$ by the sum of its entries. This normalization avoids the concentration of values at certain rows or columns of the flattening matrices which might lead to incorrect inference of the rank. As a variation we can consider the distance $\delta_{4m}(\cdot)$ to rank $4m$ matrices when we deal with mixed data with $m = 2, 3$ categories. Indeed, if we consider mixtures on $T_{A|B}$ with m categories, the corresponding flattening matrix $Flat_{A|B}$ is a sum of m rank 4 matrices. The correct quartet should be the one that gives the minimal value $s_{A|B}(F)$.

In order to have a weighting system that allows us to compare and represent the output of `Erik+2` we consider $w_{A|B}(F) = 1/s_{A|B}(F)$, and then divide by $w := w_{12|34}(F) + w_{13|24}(F) + w_{14|23}(F)$. Finally, `Erik+2` outputs the normalized three scores:

$$\texttt{Erik+2}(F) = \frac{1}{w}\left(w_{12|34}(F), w_{13|24}(F), w_{14|23}(F)\right).$$

`Erik+2` is a statistically consistent method: if $F \sim \text{Mult}\,(N; \phi_T(\theta))$ with $T = T_{12|34}$, then $\lim_{N\to\infty} \texttt{Erik+2}(F) = (1, 0, 0)$.

1.4.5 Quartet-based methods

Dealing with larger trees can have high computational cost and may become unfeasible. A common approach to reconstruct a phylogenetic tree given an alignment with a large number of taxa is to reconstruct all possible subquartets and then build the larger tree. This idea gives rise to the so-called quartet-based methods. The *quartet-based* methods (*Q-methods* for short) determine n-leaf trees given a certain reliability measure for each of the three possible topologies of every subset of 4 taxa from the original tree. This reliability is given by a system of weights.

Let F be a vector of observed relative frequencies obtained from an alignment for a set of n taxa X. Denote by $X_i = \{a, b, c, d\}$ the subsets of four elements of X and by p_i the observed relative frequencies of the taxa in X_i. The quartet-based methods described below take as input a system of weights $w_{ab|cd}(p_i), w_{ac|bd}(p_i), w_{ad|bc}(p_i)$ for each p_i and attempt to construct an n-leaf tree based on them. These methods can consider any of the

system of weights considered above. The order on which the subsets X_i are considered may affect the n-leaf tree output by Q-methods. In order to reduce the effect of this problem, Q-methods are usually applied several times, producing an n-leaf tree each time, and then consider the majority-rule consensus tree of the n-leaf trees generated:

Definition 1.4.3. Given k phylogenetic trees $T_1, \dots, T_k$ with the same set of labels at the leaves, the *majority-rule consensus tree* is the tree whose edge splits are displayed in more than half of the T_i.

It can be proved that such majority-rule consensus tree always exist (see Margush and McMorris, 1981).

Quartet Puzzling

Quartet Puzzling (QP) is a Q-method introduced by Strimmer and Haeseler (1996). It constructs an n-leaf tree from a system of weights $(w_{ab|cd}, w_{ac|bd}, w_{ad|bc})$ for each subset of 4 taxa as follows. First, it randomly fixes an order on the taxa: suppose the taxa X is $\{1, \dots, n\}$ with the natural order). Then, the algorithm starts with the set of taxa $\{1, 2, 3, 4\}$ and considers the quartet $T_{ab|cd}$ that maximizes $w_{ab|cd}$. On iterative steps, QP will be adding taxa (and therefore, nodes and edges on the tree) one by one, following the order previously determined, until the n-leaf tree is determined. At each step QP needs to add taxon i at some branch. For any choice of 3 leaves x, y, z of the partially constructed tree T_i the method adds the "penalty" $w_{xy|zi}$ to the edges in the path from x to y (as it is a measure of how bad it would be to add taxon i at the path from x to y). Finally, i is added on the least penalized branch. The same steps are repeated until all taxa are added.

It is usual to repeat this construction several times for different initial orders and then compute the majority-rule consensus tree of all these n-leaf trees.

Weight Optimization

Weight Optimization (WO) proposed by Ranwez and Gascuel (2001) follows a similar approach to QP but with some differences. The first one is that WO defines the order in which taxa will be added in a dynamic way by starting from a random subset of four taxa. At each step, the selected taxon to be added is the one that generates the greatest increase of the total weight on the tree. The other main difference is that instead of penalizing branches, WO gives branches "bonuses" and then when adding a taxon, WO chooses the branch that provides the highest increase of the total weight. We refer to Ranwez and Gascuel (2001) for more details on how these branches bonuses are given.

WO is known to reconstruct the correct tree if the input quartets are correctly weighted (see Ranwez and Gascuel, 2001). As QP, it is usual to turn this method several times for different initial quartet and then construct the majority-rule consensus tree.

Willson method

The last Q-method we recall is Willson's method (Willson, 1999). Instead of constructing a tree that maximizes the total weight at each step, the essential idea of Willson's method (briefly WIL) is the following: at each step, it attaches a new taxon in such a way that the resulting tree is the least inconsistent with the previous one. In this method the input are some *costs* weights for each quartet. The weights we have defined above are related to the reliability of each quartet: the largest the weight, the more reliable the tree. These weights w can be converted into costs by taking $-\log w$. We refer to the quoted paper for a precise explanation of the method.

As the previous methods, WIL starts from a random subset of four taxa. One usually runs the algorithm several times and then constructs the majority-rule consensus tree.

2

DISTANCE TO THE STOCHASTIC REGION OF PHYLOGENETIC VARIETIES

In this chapter we want to deal with Problem 2 presented in the Introduction. We aim to answer the explicit following question in terms of the Euclidean distance and trees of four species:

Question 2.1. *If $p \in \mathbb{R}^{4^4}$ is a distribution satisfying*

$$d\left(p, \mathcal{V}_{T_{12|34}}\right) < min\left\{d\left(p, \mathcal{V}_{T_{13|24}}\right), d\left(p, \mathcal{V}_{T_{14|23}}\right)\right\},$$

would it be possible that $d\left(p, \mathcal{V}^+_{T_{12|34}}\right) > \min\left\{d\left(p, \mathcal{V}^+_{T_{13|24}}\right), d\left(p, \mathcal{V}^+_{T_{14|23}}\right)\right\}$?

We address this problem for two special cases of interest in phylogenetics: short branches at the external edges (see Section 2.2) and long branch attraction (in Section 2.4). Both cases, short and long branches, usually lead to confusing results in phylogenetic reconstruction (particularly in relation to the long branch attraction problem, see Section 2.4). In the first case we are able to deal with the Kimura 3-parameter model and in the second case we have to restrict to the more simple Jukes-Cantor model. The reason for this restriction is that the computations get more involved in the second case and we have to use computational algebra techniques (for which is crucial to decrease the number of variables of the problem). To this end, in Section 2.3 we introduce an algorithm that computes the distance of a point to the stochastic phylogenetic regions in the JC69

case; this algorithm makes explicit use of the Euclidean distance degree (introduced in Section 1.2.3) of the phylogenetic varieties.

In this chapter we consider only the Euclidean distance. One reason for this is that the algebraic tools that we presented, such as the rank conditions on the flattening matrices, naturally deal with it. But another motivation is that the algebraic expression of the Euclidean distance permits the use of algebraic tools to derive analytical results and the use of numerical algebraic geometry to get global minima. On the other hand, the use of other distance measures such as the Hellinger distance or maximum likelihood, would not allow the use of the Fourier transform for the evolutionary models we use here, which significantly simplifies the computations in our case.

The organization of this chapter is as follows. In Section 2.1 we prove some technical results regarding the closest stochastic matrix to a given matrix. In Section 2.2 we consider the case of short external branches for the Kimura 3-parameter model and obtain analytical results. In Section 2.3 we introduce our computational approach to compute the distance to the stochastic phylogenetic regions. The results for the long branch attraction case are expanded in Section 2.4, and in Section 2.5 we provide results on simulated data that illustrate our findings.

2.1 THE CLOSEST STOCHASTIC MATRIX

Recall we write $\mathcal{H}_n$ for the hyperplane of (1.14) and $\Delta^n \subset \mathcal{H}_n$ for the standard simplex in $\mathbb{R}^n$. In this section we denote points in bold. Given a point $\mathbf{x} \in \mathbb{R}^n$, we denote by $\text{proj}_{\mathcal{H}}(\mathbf{x})$ its orthogonal projection onto $\mathcal{H}$.

Definition 2.1.1. For any matrix $M \in \mathcal{M}_n(\mathbb{R})$ we denote by $\widehat{M}$ its closest stochastic matrix in the Frobenious norm:

$$\widehat{M} = \underset{X \in \mathbb{M}_n^+}{\arg\min} \parallel M - X \parallel_F .$$

Similarly, for any point $\mathbf{x} \in \mathbb{R}^n$ we write $\widehat{\mathbf{x}}$ for its closest point in Δ^n in the Euclidean norm.

The problem of finding the nearest stochastic matrix is equivalent to finding the closest point (in Euclidean norm) in the standard simplex to every row r of the matrix, see Kreinin and Sidelnikova (2001). The uniqueness of $\widehat{r}$, and consequently of $\widehat{M}$, is guaranteed since both the objective function and the domain set are convex.

The problem of finding the closest point in the simplex Δ^n to a given point has been widely studied and there exist several algorithms to compute it. We use the Algorithm 2.1 proposed by Michelot (1986) that, given any point $\mathbf{x} \in \mathbb{R}^n$, produces the point $\widehat{\mathbf{x}} \in \Delta^n$ that minimizes $\| \mathbf{x} - \mathbf{y} \|_2$ for $\mathbf{y} \in \Delta^n$. The algorithm finishes in, at most, n steps.

Algorithm 2.1: Projection onto the simplex by Michelot (1986).

Input: $x \in \mathbb{R}^n$
while $\sum x_i \neq 1$ *or* $x \ngeq 0$ **do**
  $n_0 = \#\{x_i \neq 0\}$;
  $\lambda = \dfrac{\sum_{i=1}^n x_i - 1}{n_0}$;
  for $i = 1$ *to* n **do**
    $x_i = \max\{0, x_i - \lambda\}$;
  $\hat{x} = x$;
Output: $\hat{x} = \arg\min_{y \in \Delta^n} \| x - y \|_2$

In the following result we state some properties of the nearest points in the simplex that will be useful for determining the closest stochastic matrix. In Figure 2.1 there is an illustration of the last item.

Lemma 2.1.2. *Let $\mathbf{x} = (x_1, \dots, x_n)$ be a point in $\mathbb{R}^n$ and let $\widehat{\mathbf{x}} = (\widehat{x}_1, \dots, \widehat{x}_n)$ be its closest point in Δ^n.*

(i) $\widehat{\mathbf{x}}$ *coincides with the closest point to* $\mathrm{proj}_{\mathcal{H}}(\mathbf{x})$ *in* Δ^n, $\widehat{\mathrm{proj}_{\mathcal{H}}(\mathbf{x})}$.

(ii) If $\mathbf{x} \in \mathcal{H}$ *and* $x_i \leq 0$ *for some i, then* $\widehat{x}_i = 0$.

(iii) Let $\mathbf{y}$ *be a point obtained by a permutation of the coordinates of* $\mathbf{x}$, *i.e.* $\mathbf{y} = P\mathbf{x}$ *for some permutation matrix P. Then* $\widehat{\mathbf{y}} = P\widehat{\mathbf{x}}$.

(iv) If $x_i = x_j$ *for some* $i, j = 1, \dots, n$ *then* $\widehat{x}_i = \widehat{x}_j$.

(v) $\widehat{\mathbf{x}}$ *coincides with* $\mathbf{p}_i = (0, \dots, \underset{\widehat{i}}{1}, \dots, 0)$ *if and only if* $x_i - x_j \geq 1 \; \forall j \neq i$.

Proof. The proofs of items (i) and (ii) can be found in the paper by Michelot (1986) and are the basis of the Algorithm 2.1.
(iii) Follows from the fact that P is a permutation matrix and hence it is orthogonal.
(iv) Is a direct consequence of (iii).
(v) Using (i) and (ii) we can assume that $\sum_i x_i = 1$, i.e., $\mathbf{x}$ belongs to the affine hyperplane $\mathcal{H}$. By symmetry, it is enough to prove the result for $\mathbf{p}_1$, that is, we prove that $\widehat{\mathbf{x}} = \mathbf{p}_1$

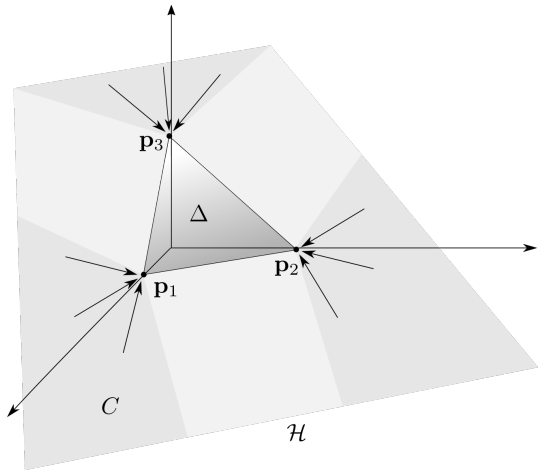


FIGURE 2.1: The hyperplane $\mathcal{H}$ and the standard simplex Δ^n are represented in the case $n = 3$. The 2-dimensional cone C given by the inequalities $x_1 - x_2 \geq 1$, $x_1 - x_3 \geq 1$ corresponds to all the points in $\mathcal{H}$ whose projection on the simplex is $\mathbf{p}_1$ (see (v) in Lemma 2.1.2).

if and only if $x_1 - x_j \geq 1$ for all $j \neq 1$. Firstly we show that if $x_1 - x_j \geq 1$, $j \neq 1$, then necessarily $\widehat{\mathbf{x}} = \mathbf{p}_1$. Indeed, if $\mathbf{q} = (c_1, \dots, c_n) \in \Delta^n$, we have that

$$\begin{aligned} d(\mathbf{x}, \mathbf{q})^2 &= \sum_{j=1}^{n} x_j^2 + \left(\sum_{j=1}^{n} c_j^2 - 2 \sum_{j=1}^{n} c_j x_j \right), \\ d(\mathbf{x}, \mathbf{p}_1)^2 &= \sum_{j=1}^{n} x_j^2 + (1 - 2x_1). \end{aligned} \tag{2.1}$$

Now, because of the assumption $x_1 - x_j \geq 1$ and $\sum_j c_j = 1$, we have that

$$\sum_{j=1}^{n} c_j x_j \leq \sum_{j=1}^{n} c_j x_1 - \left(\sum_{j=2}^{n} c_j \right) = x_1 + (c_1 - 1).$$

In particular,

$$\sum_{j=1}^{n} c_j^2 - 2 \sum_{j=1}^{n} c_j x_j \geq \sum_{j=1}^{n} c_j^2 - 2(x_1 + c_1 - 1) = (c_1 - 1)^2 + \sum_{j=2}^{n} c_j^2 + (1 - 2\,x_1) \geq 1 - 2\,x_1.$$

Comparing this with (2.1), it follows that $d(\mathbf{x}, q) \geq d(\mathbf{x}, \mathbf{p}_1)$ for any $q \in \Delta^n$, so $\mathbf{p}_1 = \widehat{\mathbf{x}}$. Conversely, assume that $\mathbf{x} \in \mathcal{H}$ is such that $x_1 - x_i < 1$ for some $i \geq 2$. We will show that there exists some $\mathbf{q}$ in the edge $\mathbf{p}_1\mathbf{p}_i$ such that $d(\mathbf{x}, \mathbf{q}) < d(\mathbf{x}, \mathbf{p}_1)$ so that $\mathbf{p}_1$ cannot be the

the closest point to $\mathbf{x}$ in the simplex Δ^n. Consider $\mathbf{q} = a\mathbf{p}_1 + b\mathbf{p}_i$ with $a, b \geq 0$, $a + b = 1$. As above, we have that

$$d(\mathbf{x}, \mathbf{q})^2 = \sum_{j=1}^{n} x_j^2 + a^2 - 2\,a\,x_1 + b^2 - 2\,b\,x_i.$$

We claim that if we take $0 < b < 1 + x_i - x_1$, then the point $\mathbf{q}$ satisfies the inequality between distances above. Indeed, we need to verify that

$$a^2 - 2\,a\,x_1 + b^2 - 2\,b\,x_i < 1 - 2x_1 = (a+b)^2 - 2x_1 = a^2 + b^2 + 2ab - 2x_1.$$

Using that $a = 1 - b$, this is equivalent to the inequality $b(b - 1 + x_1 - x_i) < 0$, which is satisfied by our choice of b. □

Remark 2.1.3. If the rows of a matrix M are the result of some permutation applied to the first row, the previous lemma shows that $\widehat{M}$ will preserve the same identities between entries as the matrix M. In particular, if M is a transition matrix of a group-based model (see Remark 1.1.29), then $\widehat{M}$ will remain in the same model. For example, consider the non-stochastic K81 matrix (see Definition 1.1.27)

$$M = \begin{pmatrix} 0.91 & 0.08 & -0.09 & 0.1 \\ 0.08 & 0.91 & 0.1 & -0.09 \\ -0.09 & 0.1 & 0.91 & 0.08 \\ 0.1 & -0.09 & 0.08 & 0.91 \end{pmatrix}.$$

Applying Algorithm 2.1 to $x = (0.91, 0.08, -0.09, 0.1)$ one gets,

`Step 1:` $\lambda_1 = 0$, $\hat{x}_1 = (0.91, 0.08, 0, 0.1)$,

`Step 2:` $\lambda_2 = 0.03$, $\hat{x}_2 = (0.88, 0.05, 0, 0.07)$.

`Output:` $\hat{x} = (0.88, 0.05, 0, 0.07)$

Then, the nearest stochastic matrix to M is

$$\widehat{M} = \begin{pmatrix} 0.88 & 0.05 & 0 & 0.07 \\ 0.05 & 0.88 & 0.07 & 0 \\ 0 & 0.07 & 0.88 & 0.05 \\ 0.07 & 0 & 0.05 & 0.88 \end{pmatrix},$$

which preserves the same identities between entries and so, it is a stochastic K81 matrix.

Lemma 2.1.4. *Let M be a JC69 matrix. Then M is stochastic if and only if its eigenvalues lie in $[-1/3, 1]$.*

Proof. Let M be a JC69 matrix, that is, M as in (1.7) with $c = d = b$, $a = 1 - 3b$. Then, M is stochastic if and only if $b \geq 0$ and $a = 1 - 3b \geq 0$, which is equivalent to $b \in [0, 1/3]$. As the eigenvalues of M are $m_{\mathtt{A}} = 1$ and $m_{\mathtt{C}} = m_{\mathtt{G}} = m_{\mathtt{T}} = 1 - 4b$ (see Lemma 1.1.28), we get that M is stochastic if and only if the eigenvalue $1 - 4b$ lies in $[-1/3, 1]$. □

Lemma 2.1.5. *Let M be a non-stochastic JC69 matrix. Then $\widehat{M}$ is either the identity matrix or the matrix*

$$\begin{pmatrix} 0 & 1/3 & 1/3 & 1/3 \\ 1/3 & 0 & 1/3 & 1/3 \\ 1/3 & 1/3 & 0 & 1/3 \\ 1/3 & 1/3 & 1/3 & 0 \end{pmatrix}.$$

Proof. Let M be a JC69 matrix with off-diagonal entries equal to b and diagonal entries equal to $a = 1 - 3b$. Then it is not stochastic if either $b < 0$ or $a < 0$. Let $v = (a, b, b, b)$ be the first row of M. Then by Lemma 2.1.2 (iv), its projection onto the simplex Δ^3 is of the type $\widehat{v} = (\widehat{a}, \widehat{b}, \widehat{b}, \widehat{b})$. The following argument is valid for each row due to Lemma 2.1.2 (iii).

If $b < 0$ then, by Lemma 2.1.2 (ii), $\widehat{b}$ equals zero and $\widehat{a}$ has to be equal to 1 since the coordinates of $\widehat{v}$ sum to 1. Therefore $\widehat{M}$ is the 4×4 identity matrix.

If $a < 0$ then $\widehat{a} = 0$ and since $3\widehat{b} = 1$, $\widehat{b} = \frac{1}{3}$. Therefore $\widehat{M}$ is a matrix with 0 in the diagonal and $\frac{1}{3}$ at the non-diagonal entries. □

For later use, we close this section by stating a characterization of those K81 matrices M for which $\widehat{M}$ is a permutation matrix.

Lemma 2.1.6. *Let M be a K81 matrix and denote by (a_1, a_2, a_3, a_4) its first row. Then $\widehat{M}$ is a permutation matrix if and only if there is some $i \in \{1, \dots, 4\}$ such that*

$$a_i - a_j \geq 1 \text{ for all } j \neq i.$$

Proof. This is an immediate consequence of Lemma 2.1.2 (v). □

2.2 THE CASE OF SHORT EXTERNAL BRANCHES

In this section we study evolutionary processes where substitutions at the external edges are unusual, so that probabilities of substitution of nucleotides in the corresponding transition matrices are small. This translates to matrices close to the identity at the external edges and *short branch lengths*, as explained in Section 1.1.4.

We use the results of Section 2.1 with $n = 4^4$ and we stick to the K81 model. Given $p \in \mathbb{R}^{4^4}$, let p_T^+ be a point in the stochastic region $\mathcal{V}_T^+$ that minimizes the distance to p, i.e.

$$d(p, p_T^+) = d(p, \mathcal{V}_T^+).$$

The following result shows that a point p arising from a tree T with the identity matrix attached at the external edges, that is $p = \phi_T(Id, Id, Id, Id, M)$, is always closer to the stochastic region of T than to any other stochastic phylogenetic region. See Figure 2.2 for an illustration of this result.

Proposition 2.2.1. *Assume that $p = \phi_T(Id, Id, Id, Id, M)$ where M is a non-stochastic K81 matrix and T is any quartet. Then,*

(a) The point p_T^+ is equal to $\phi_T(Id, Id, Id, Id, \widehat{M})$. Moreover, p_T^+ coincides with the point that minimizes the distance to the standard simplex $\Delta := \Delta^{4^4} \subset \mathbb{R}^{4^4}$. In particular, the point p_T^+ is unique.

(b) If $T' \neq T$ is another tree in $\mathcal{T}_4$, then $d(p, \mathcal{V}_{T'}^+) \geq d(p, \mathcal{V}_T^+)$.

(c) The following are equivalent:

(i) equality holds in (b);

(ii) $p_T^+ \in \mathcal{V}_T^+ \cap \mathcal{V}_{T'}^+$;

(iii) the matrix $\widehat{M}$ is a permutation matrix.

Proof. We assume that $T = T_{12|34}$, but the proof is analogous for the other trees. We define $\widehat{p}$ to be the closest point to p in Δ (which is a convex set), see Lemma 2.1.2. First of all, as $\mathcal{V}_T^+ \subset \Delta$, we have that

$$d(p, \mathcal{V}_T^+) = \min_{q \in \mathcal{V}_T^+} d(p, q) \geq \min_{q \in \Delta} d(p, q) = d(p, \widehat{p}). \tag{2.2}$$

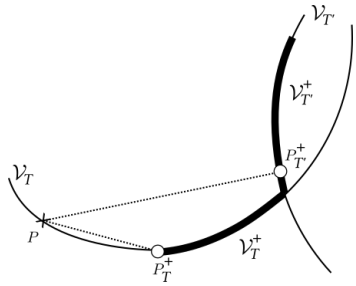


FIGURE 2.2: For two different quartet trees T, T', the phylogenetic varieties $\mathcal{V}_T$ and $\mathcal{V}_{T'}$ are represented as curves, with the intersection being reduced to only one point. The stochastic regions are represented with thick stroke. The point $p = \phi(Id, Id, Id, Id, M)$ with M non-stochastic lies in $\mathcal{V}_T$ but not in the stochastic region $\mathcal{V}_T^+$. The points p_T^+ and $p_{T'}^+$ represent points that minimize the distance from p to $\mathcal{V}_T^+$ and $\mathcal{V}_{T'}^+$, respectively. The figure illustrates that $d(p, \mathcal{V}_T^+) \leq d(p; \mathcal{V}_{T'}^+)$ (see (b) of Proposition 2.2.1).

We now show that $\widehat{p} \in \mathcal{V}_T^+$. Since the transition matrices at the exterior edges of T are the identity matrix, the coordinates of p are

$$p_{ijkl} = \begin{cases} \frac{1}{4}(M)_{ik} & \text{if } i = j \text{ and } k = l \\ 0 & \text{otherwise.} \end{cases}$$

Since M is a K81 matrix, the non-zero coordinates of p only take 4 different values. Moreover, because of Lemma 2.1.2 (ii) and (iv), we can write the coordinates of $\widehat{p}$ as

$$\widehat{p}_{ijkl} = \begin{cases} b_{ik} & \text{if } i = j \text{ and } k = l \\ 0 & \text{otherwise} \end{cases}$$

for some values b_{ik} satisfying the identities of a K81 matrix (see (1.7)). Since $\widehat{p}$ belongs to the simplex, we have that $\sum_{i,k} b_{ik} = 1$. It follows that the matrix

$$4 \begin{pmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ b_{31} & b_{32} & b_{33} & b_{34} \\ b_{41} & b_{42} & b_{43} & b_{44} \end{pmatrix}$$

is a K81 stochastic matrix. Actually, this matrix is just $\widehat{M}$ (again by Lemma 2.1.2), and so, $\widehat{p} = \phi_T(Id, Id, Id, Id, \widehat{M})$. In particular, $\widehat{p} \in \mathcal{V}_T^+$. Since p_T^+ minimizes the distance from p to the variety $\mathcal{V}_T^+$, we have $d(p, \widehat{p}) \geq d(p, p_T^+)$. Because of (2.2), equality holds.

Moreover, from the uniqueness of the point minimizing the distance to Δ, it follows that $p_T^+ = \widehat{p}$. This concludes the proof of (a).

(b) For any tree topology T', we have that $\mathcal{V}_{T'}^+ \subset \Delta$. It follows that $d(p, \widehat{p}) \leq d(p, p_{T'}^+)$. Since $\widehat{p} = p_T^+$, we infer that $d(p, p_T^+) \leq d(p, p_{T'}^+)$ for any $T' \neq T$.

(c) Now, we proceed to characterize when the equality holds in (b).
(i) $\Leftrightarrow$ (ii). It is clear that if $p_T^+ \in \mathcal{V}_{T'}^+$, then $d(p, \mathcal{V}_T^+) = d(p, p_T^+) \geq d(p, \mathcal{V}_{T'}^+)$. Together with the inequality in (b), this proves that (ii) implies (i). Conversely, if the equality holds, then $d(p, p_{T'}^+) = d(p, \Delta)$. Because of the uniqueness of the point that minimizes the distance to Δ, it follows that $p_{T'}^+ = \widehat{p}$, and we have already seen that $\widehat{p} \in \mathcal{V}_T^+$. Therefore, $p_T^+ \in \mathcal{V}_T^+ \cap \mathcal{V}_{T'}^+$.

(ii) $\Leftrightarrow$ (iii). It only remains to see that $p_{T'}^+ = p_T^+$ (i.e. $p_T^+ \in \mathcal{V}_T^+ \cap \mathcal{V}_{T'}^+$) if and only if M is a permutation matrix. Assume that $\widehat{p} \in \mathcal{V}_{T'}^+$ then if $A'|B'$ is a compatible split for T' then $flatt_{A'|B'}(\widehat{p})$ has rank less or equal to 4 (see Theorem 1.3.24). On the other hand, as $\widehat{p} = \phi_T(Id, \dots, Id, \widehat{M})$, $flatt_{T'}(p)$ is a diagonal matrix whose diagonal is formed by the 16 entries of $\widehat{M}$ multiplied by a constant (see Lemma 1.3.23). The only way this matrix has rank ≤ 4 is by imposing the vanishing of 12 entries. Since M is a K81 stochastic matrix, $\widehat{M}$ has to be a permutation matrix. Conversely, if $\widehat{M}$ is a permutation matrix, then the corresponding point $\widehat{p} = \phi_T(Id, \dots, Id, \widehat{M})$ lies in variety $\mathcal{V}_{T'}^+$ for every $T' \in \mathcal{T}_4$. □

Remark 2.2.2. Note that p_T^+ arises from the parameters $\{Id, Id, Id, Id, \widehat{M}\}$ but also from the parameters obtained by a label swapping (see Theorem 1.3.6 and the discussion below).

In the following theorem we prove that if $\widehat{M}$ is not a permutation matrix and p is close enough to the point $p_0 = \phi_T(Id, Id, Id, Id, M)$ then p will remain closer to the stochastic region $\mathcal{V}_T^+$ than to the stochastic region $\mathcal{V}_{T'}^+$ for $T' \neq T$ (see Figure 2.3 for an illustration). We need to exclude the case $d(p, \mathcal{V}_T^+) = d(p, \mathcal{V}_{T'}^+)$ (case (c) of Proposition 2.2.1) if we want strict inequality.

Theorem 2.2.3. *Let M be a K81 non-stochastic matrix for which $\widehat{M}$ is not a permutation matrix (see Lemma 2.1.6 for a characterization). Let $p_0 = \phi_T(Id, Id, Id, Id, M)$, $T' \in \mathcal{T}_4 \setminus \{T\}$, and let $p \in \mathbb{R}^{4^4}$ be a point such that*

$$d(p, p_0) < \frac{d(p_0, \mathcal{V}_{T'}^+) - d(p_0, \mathcal{V}_T^+)}{2}$$

(this is satisfied if p is close enough to p_0). Then $d(p, \mathcal{V}_T^+) < d(p, \mathcal{V}_{T'}^+)$.

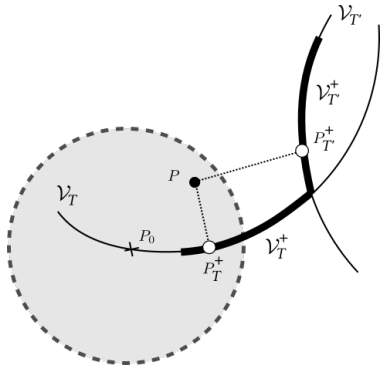


FIGURE 2.3: The point p_0 (see Theorem 2.2.3) lies in $\mathcal{V}_T$ but not in $\mathcal{V}_T^+$. As long as a point p lies close to p_0, namely $d(p, p_0) < (d(p_0, \mathcal{V}_{T'}^+) - d(p_0, \mathcal{V}_T^+))/2$, it will remain closer to the stochastic region $\mathcal{V}_T^+$ than to the stochastic region $\mathcal{V}_{T'}^+$ for $T' \neq T$.

Proof. We first define the function $f(q) = d(q, V_{T'}^+) - d(q, V_T^+)$. By hypothesis, $\widehat{M}$ is not a permutation matrix and by Proposition 2.2.1, we have that $f(p_0) > 0$. We want to show that $f(p) > 0$ if $d(p, p_0) < f(p_0)/2$. Clearly, we are done if $f(p) \geq f(p_0)$, so we assume that $f(p) < f(p_0)$. From the triangle inequality we have $|d(p, \mathcal{W}) - d(p_0, \mathcal{W})| \leq d(p, p_0)$, for any $\mathcal{W} \subset \mathbb{R}^N$. Then, we obtain

$$\begin{aligned} |f(p) - f(p_0)| &= |d(p, \mathcal{V}_{T'}^+) - d(p_0, \mathcal{V}_{T'}^+) - \big(d(p, \mathcal{V}_T^+) - d(p_0, \mathcal{V}_T^+)\big)| \leq \\ &\leq |d(p, \mathcal{V}_{T'}^+) - d(p_0, \mathcal{V}_{T'}^+)| + |d(p, \mathcal{V}_T^+) - d(p_0, \mathcal{V}_T^+)| \\ &\leq 2\, d(p, p_0) < f(p_0). \end{aligned}$$

Therefore, $f(p) = (f(p) - f(p_0)) + f(p_0) = -|f(p) - f(p_0)| + f(p_0) > 0$. This concludes the proof. □

Example 2.2.4. The matrix M of Remark 2.1.3 satisfies the hypothesis of Theorem 2.2.3.

2.3 COMPUTING THE CLOSEST POINT TO A STOCHASTIC PHYLOGENETIC REGION

Although in the last section we were able to answer our questions analytically, this approach seems unfeasible if we want to tackle more general problems. In this section, in order to find the distance from a point to a stochastic phylogenetic variety we propose an alternative approach based on numerical algebraic geometry. Our goal is to find all critical points of the distance function to a phylogenetic variety in the interior and at the boundary of the stochastic region. Among the set of critical points we pick the ones that

minimize the distance. Similar approaches, where computational and numerical algebraic geometry are applied to phylogenetics studies, can be found in the works by Gross et al. (2016) and by Kosta and Kubjas (2019).

In this section we assume the JC69 model and we parametrize each transition matrix by its eigenvalue different from 1 (see Lemma 1.1.28). We use the following notation.

Notation 2.3.1. Consider a quartet T with JC69 matrices M_i attached at the edges of the tree. Denote by x_i the eigenvalue of M_i of multiplicity three different from 1 (see Lemma 1.1.28) and we denote by φ_T the parameterization of the phylogenetic varieties on the Fourier parameters to Fourier coordinates (see Section 1.3.2),

$$\begin{array}{rcl} \varphi_T : \mathbb{R}^5 & \longrightarrow & \mathbb{R}^{4^4} \\ \mathbf{x} = (x_1, x_2, x_3, x_4, x_5) & \longmapsto & \bar{p} = \varphi_T(x_1, x_2, x_3, x_4, x_5). \end{array}$$

The parametrization φ_T can be computed by adapting Theorem 1.3.19 to this model as in Example 1.3.20.

Recall that, by Lemma 2.1.4, $\varphi_T(x_1, \dots, x_5)$ is a point in the stochastic region if and only if $x_i \in [-1/3, 1]$, $i = 1, \dots, 5$. We denote by

$$\mathcal{D} := [-1/3, 1]^5 \tag{2.3}$$

the region of stochastic parameters.

Given a point $p \in \mathbb{R}^{4^4}$, we denote by $f_T(x_1, \dots, x_5)$ the square of the Euclidean distance function from the point $\varphi_T(x_1, \dots, x_5)$ to p:

$$f_T(x_1, \dots, x_5) = d(p, \varphi_T(x_1, \dots, x_5))^2.$$

Recall that we can use Fourier coordinates to compute distances between points because of Remark 1.3.18. We view f_T as a function from $\mathbb{R}^5$ to $\mathbb{R}$, or when convenient from $\mathbb{C}^5$ to $\mathbb{C}$, and our goal is to find a minimum of f_T at $\mathcal{D}$.

In order to compute the number of critical points of this function we use the Euclidean distance degree, as introduced in Section 1.2.3. To this end we need to know the set of singular points of the corresponding variety.

Under the Jukes-Cantor model, the singular points of the varieties $\mathcal{V}_T$ are those that are the image of some null parameter (see Casanellas and Fernández-Sánchez (2008) and Casanellas, Fernández-Sánchez, and Michałek (2015) for details). In other words, $\varphi_T(x_1, \dots, x_5)$ is a singular point of the variety if and only if $x_i = 0$ for some i.

Hence, we can compute the number of critical points of our function f_T in the preimage of the smooth part of the variety as the degree of saturation ideal $I : (x_1 \cdot \ldots \cdot x_5)^\infty$, where I is generated by the partial derivatives of f_T. Using this, we have computed the the EDdegree of $\mathcal{V}_T$ for several random points using the Code A.1 in Appendix A. The algorithm has been implemented in `Magma` (Bosma, Cannon, and Playoust, 1997) and we have obtained:

Lemma 2.3.2. *If $\mathcal{V}_T$ is the phylogenetic variety corresponding to a quartet evolving under the JC69 model, then the EDdegree of $\mathcal{V}_T$ is* 290.

To identify the critical points of our constrained problem we use the *KKT conditions of first order for local minimums*. We recall this procedure below.

Karush-Kuhn-Tucker conditions (KKT)

If $f, g_i : \mathbb{R}^l \longrightarrow \mathbb{R}$ are $\mathcal{C}^\infty$ functions, for $i = 1, \ldots, n$, we consider the following optimization problem:

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & f(\mathbf{x}) \\ \text{subject to} \quad & g_i(\mathbf{x}) \leq 0, \ i = 1, \ldots, n. \end{aligned}$$

If a point $\mathbf{x}^*$ that satisfies $g_i(\mathbf{x}^*) \leq 0 \ \forall i = 1, \ldots, m$ is a local optimum of the problem, then there exist some constants $\mu = (\mu_1, \ldots, \mu_n)$ (called *KKT multipliers*) such that $\mathbf{x}^*$ and μ satisfy

(i) $-\nabla f(\mathbf{x}^*) = \sum_{i=1}^n \mu_i \nabla g_i(\mathbf{x}^*)$,

(ii) $\mu_i \geq 0 \ \forall i = 1, \ldots, n$,

(iii) $\mu_i g_i(\mathbf{x}^*) = 0 \ \forall i = 1, \ldots, n$.

According to these conditions, the algorithm to find the closest points in $\mathcal{V}_T$ to a given point falls naturally into two parts:

- find the 290 critical points of the objective function f_T over all $\mathbb{C}^5$ and keep only those points with *real* coordinates that *lie* in $\mathcal{D}$.
- check the optimal points at the boundary of $\mathcal{D}$.

Write

$$g_{1,i}(\mathbf{x}) := x_i - 1 \qquad g_{2,i}(\mathbf{x}) := -x_i - 1/3$$

and consider the inequalities $g_{1,i}(\mathbf{x}) \leq 0$ and $g_{2,i}(\mathbf{x}) \leq 0$ defining the region $\mathcal{D}$. For each $i = 1, \dots, 5$ and $l = 1, 2$, write

$$S_{l,i} = \{\mathbf{x} = (x_1, \dots, x_5) \mid g_{l,i} = 0\}.$$

To find the critical points at the boundary of $\mathcal{D}$ we restrict f_T to its faces, that is the subsets $S_{\iota_1,\iota_2} := \left(\bigcap_{i\in\iota_1} S_{1,i}\right) \bigcap \left(\bigcap_{j\in\iota_2} S_{2,j}\right)$ for some disjoint $\iota_1, \iota_2 \subseteq \{1, \dots, 5\}$, and find critical points there. Then we only consider those that satisfy $g_{1,i}(\mathbf{x}) \leq 0$ and $g_{2,i}(\mathbf{x}) \leq 0$.

We use homotopy continuation methods to solve the different polynomial systems $\nabla f_T = 0$, $\nabla f_T \mid_S = 0$. All computations have been done with the package `PHCpack.m2` (see Verschelde, 1999 and Gross, Petrović, and Verschelde, 2013) which turned out to be the only numerical package capable to find these 290 points of $I : (x_1 \cdot \ldots \cdot x_5)^\infty$. `Macaulay2` (Grayson and Stillman, 2009) has been used to implement the main core of the algorithm while some previous computations have been performed in `Magma` (Bosma, Cannon, and Playoust, 1997). The whole code can be found in Code A.2 and A.3 in Appendix A.

Algorithm 2.2: The closest point in the stochastic phylogenetic region $\mathcal{V}_T^+$

Input: A point $p \in \mathbb{R}^{4^4}$ and a quartet T.

Set f_T;
Set $\mathcal{J} := \left(\partial_{x_1}(f_T), \partial_{x_2}(f_T), \partial_{x_3}(f_T), \partial_{x_4}(f_T), \partial_{x_5}(f_T)\right)$;
$\mathcal{L} := \{\}$; `// Empty list of valid critical points`
$d := degree\left(I : (x_1 \cdots x_5)^\infty\right)$; `// d=290`
Find the d 0-dimensional solutions of ∇f_T=0 ;
foreach *solution* x **do**
  if $\mathbf{x} \in \mathbb{R}^5$ *and* $g_{l,i}(\mathbf{x}) \leq 0\ \forall l, i$ **then**
    Add $\mathbf{x}$ to $\mathcal{L}$;
foreach *disjoint subsets* $\iota_1, \iota_2 \subseteq \{1, \dots, 5\}$ **do**
  $S := \left(\cap_{i\in\iota_1} S_{1,i}\right) \cap \left(\cap_{j\in\iota_2} S_{2,j}\right)$;
  Find the solutions of $\nabla(f_T)_{|S} = 0$;
  if $\mathbf{x} \in \mathbb{R}^5$ *and* $g_{l,i}(\mathbf{x}) \leq 0\ \forall l, i$ **then**
    Add $\mathbf{x}$ to $\mathcal{L}$;
Evaluate f_T at each $\mathbf{x} \in \mathcal{L}$ and return the point $\mathbf{x}^*$ with minimum $f_T(\mathbf{x}^*)$;

Output: Parameters $x_1^*, x_2^*, x_3^*, x_4^*, x_5^*$ such that $p_T^+ := \varphi_T(x_1^*, \dots, x_5^*) \in \mathcal{V}_T^+$ and $d(p, \mathcal{V}_T^+) = d(p, p_T^+)$.

This approach is followed in the next two sections. In Section 2.4 we use this algorithm to compute the distance from a (general) point $p \in \mathcal{V}_T$ (but not in $\mathcal{V}_T^+$) to the stochastic region $\mathcal{V}_{T'}^+$, $T' \in \mathcal{T}_4$. In Section 2.5 we use this approach to compute the distance of random points to $\mathcal{V}_T^+$.

2.4 THE LONG BRANCH ATTRACTION CASE

In this section we study the distance of p to the stochastic phylogenetic regions to give an answer to the Question 2.1 for the long branch attraction case under the Jukes Cantor model.

To this end, we keep the notation introduced in Section 2.3. Consider the tree in Figure 2.4, with a non-stochastic matrix M at the interior edge, a stochastic transition matrix K at the other edges adjacent to leaves 1 and 3, and the identity matrix Id at the remaining edges. Assume K and M are Jukes-Cantor matrices and the distribution at the root is uniform. Let k (respectively m) be the eigenvalue of K (resp. of M) different from 1. We take K stochastic, this is $k \in [-1/3, 1]$ (see Lemma 2.1.4) and we assume $m > 1$ to make M non-stochastic (the other possibility would be that $m < -1/3$, but this leads to a biologically unrealistic situation). Let $p := \varphi_{12|34}(k, 1, k, 1, m)$ be the image of these parameters written in Fourier coordinates.

Given $T \in \mathcal{T}_4$, we want to find the closest point to p in $\mathcal{V}_T^+$, that is, to find $(x_1, \dots, x_5) \in \mathcal{D}$ such that $d\left(p, \mathcal{V}_T^+\right) = d\left(p, \varphi_T(x_1, x_2, x_3, x_4, x_5)\right)$.

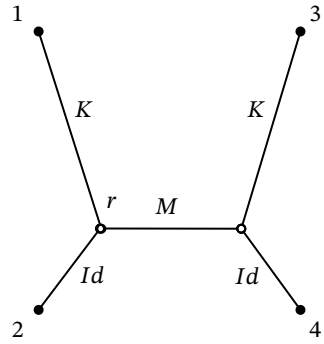


FIGURE 2.4: Phylogenetic tree $T_{12|34}$ with transition matrices giving rise to the point $P = \varphi_{12|34}(k, 1, k, 1, m)$.

Using the notation of Section 2.3, this problem can be translated into the following optimization problem:

Problem 2.4.1. Set $p = \varphi_{12|34}(k,1,k,1,m)$, then we need to

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & f_T(\mathbf{x}) := d\left(p, \varphi_T(x_1,x_2,x_3,x_4,x_5)\right)^2 \\ \text{subject to} \quad & g_{1,i}(\mathbf{x}) \leq 0,\ i = 1,\dots,5, \\ & g_{2,i}(\mathbf{x}) \leq 0,\ i = 1,\dots,5. \end{aligned}$$

where $g_{1,i}(\mathbf{x}) = x_i - 1$ and $g_{2,i}(\mathbf{x}) = -x_i - \frac{1}{3}$.

2.4.1 Local minimum

For $T = T_{12|34}$, the Euclidean distance (in Fourier coordinates) from the point $p = \varphi_{12|34}(k,1,k,1,m) \in \mathcal{V}_{12|34}$ to a point $\varphi_{12|34}(x_1,x_2,x_3,x_4,x_5) \in \mathcal{V}_{T_{12|34}}$ is proportional to the square root of the following function:

$$\begin{aligned} f_{12|34}(x_1,x_2,x_3,x_4,x_5) := {} & 12\left(x_1x_2x_3x_4x_5 - k^2m\right)^2 + 9\left(x_1x_2x_3x_4 - k^2\right)^2 \\ & + 6\left(x_1x_2x_3x_5 - k^2m\right)^2 + 6\left(x_1x_2x_4x_5 - km\right)^2 \\ & + 6\left(x_1x_3x_4x_5 - k^2m\right)^2 + 6\left(x_2x_3x_4x_5 - km\right)^2 \\ & + 3\left(x_1x_3x_5 - k^2m\right)^2 + 3\left(x_2x_3x_5 - km\right)^2 + 3\left(x_1x_4x_5 - km\right)^2 \\ & + 3\left(x_2x_4x_5 - m\right)^2 + 3\left(x_1x_2 - k\right)^2 + 3\left(x_3x_4 - k\right)^2. \end{aligned}$$

An initial numerical approach suggests a candidate $\mathbf{x}^*$ to be a minimum of the optimization problem 2.4.1 when $T = T_{12|34}$ (and also for the other trees, as we will see later). We proceed to introduce this point. Define

$$\omega := \frac{4}{9} + \frac{11}{27\sqrt[3]{\frac{69+16\sqrt{3}}{243}}} + \sqrt[3]{\frac{69+16\sqrt{3}}{243}} \approx 1.734$$

and the intervals $I := \left[-\frac{1}{3}, 1\right]$ and $\Omega := (1,\omega]$. Then, given $(k,m) \in I \times \Omega$, we denote $\mathbf{x}^* \in \mathbb{R}^5$ the following point:

$$\mathbf{x}^* = \begin{cases} \left(\tilde{x}(k,m), 1, \tilde{x}(k,m), 1, 1\right) & \text{if } \tilde{x}(k,m) < 1; \\ (1,1,1,1,1) & \text{otherwise,} \end{cases} \tag{2.4}$$

where

$$\tilde{x}(k,m) = \frac{3k^2(3m+1)-4}{36\gamma(k,m)} + \gamma(k,m), \tag{2.5}$$

$\gamma(k,m) = \sqrt[3]{\frac{1}{24}k(3m+1) + \frac{1}{216}\sqrt{\alpha(k,m)}}$ and $\alpha(k,m)$ is a positive value given by

$$\begin{aligned}\alpha(k,m) = &-729k^6m^3 - 27k^6 + 108k^4 - 243\left(3k^6 - 4k^4 - 3k^2\right)m^2 \\ &- 63k^2 - 27\left(9k^6 - 24k^4 - 2k^2\right)m + 64.\end{aligned}$$

Some comments are in order. We restrict to the cases where $(k,m) \in I \times \Omega$ since we want K to be stochastic and M to be non-stochastic but not far from the identity. Moreover, we also choose this region for computational reasons, since we will be able to prove that $\tilde{x}(k,m)$ is real and continuous in $I \times \Omega$. As the parameter of $\mathbf{x}^*$ corresponding to the interior edge is 1, $\varphi_T(\mathbf{x}^*)$ belongs to the intersection of the tree phylogenetic varieties $\mathcal{V}_{T_{12|34}} \cap \mathcal{V}_{T_{13|24}} \cap \mathcal{V}_{T_{14|23}}$ (see also Lemma 2.2.1). For that reason it is natural to ask whether $\mathbf{x}^*$ is also a local minimum of the optimization problem 2.4.1 for $T = T_{13|24}$ or $T = T_{14|23}$. In Theorem 2.4.3 we will prove that $\mathbf{x}^*$ is a local minimum of the optimization problem 2.4.1 for any $T \in \mathcal{T}_4$ however, first we prove that it is well defined.

Proposition 2.4.2. *Given $(k,m) \in I \times \Omega$ and $T \in \mathcal{T}_4$, the function $x \to f_T(x,1,x,1,1)$ has a unique critical point, which is given by the expression $\tilde{x}(k,m)$ defined in (2.5). Moreover, this expression defines a continuous function on $I \times \Omega$.*

Proof. Straightforward computations show that

$$\begin{aligned}&f_{12|34}(x,1,x,1,1) = f_{13|24}(x,1,x,1,1) = f_{14|24}(x,1,x,1,1) = \\ &27\left(x^2 - k^2m\right)^2 + 9\left(x^2 - k^2\right)^2 + 18\left(x - km\right)^2 + 6\left(x - k\right)^2 + 3\left(1 - m\right)^2\end{aligned}$$

and that the only real critical point of this function, when $(k,m) \in I \times \Omega$, is the point $\tilde{x}(k,m)$ given by expression (2.5). In order to prove that $\tilde{x}$ is a continuous real function on $I \times \Omega$, we first prove that $\gamma(k,m)$ is real and then that it does not vanish.

In order to see that $\gamma(k,m)$ is real we prove the $\alpha(k,m) \geq 0$, for all $(k,m) \in I \times \Omega$. Consider $\alpha_m(k) := \alpha(k,m)$ as a function of k, i.e. suppose m is fixed.

$$\alpha_m(k) = \underbrace{\left(-729m^3 - 729m^2 - 243m - 27\right)}_{a(m)}k^6 + \underbrace{\left(972m^2 + 648m + 108\right)}_{b(m)}k^4 + \underbrace{\left(729m^2 + 54m - 63\right)}_{c(m)}k^2 + \underbrace{64}_{d}.$$

Note that $\alpha(k,m)$ is an even function of k (i.e. $\alpha_m(k) = \alpha_m(-k)$). This function has a local minimum at $k = 0$ since

$$\begin{cases} \alpha_m(k) = a(m)k^6 + b(m)k^4 + c(m)k^2 + d & \text{and} \quad \alpha_m(0) = d = 64 > 0, \\ \alpha_m'(k) = 6a(m)k^5 + 4b(m)k^3 + 2c(m)k & \text{and} \quad \alpha_m'(0) = 0, \\ \alpha_m''(k) = 30a(m)k^4 + 12b(m)k^2 + 2c(m) & \text{and} \quad \alpha_m''(0) = 2\left(729m^2 + 54m - 63\right), \end{cases}$$

and $\alpha_m''(k)$ is strictly positive for $m > 1$. Hence, $\alpha_m(k)$ is an even polynomial of degree 6 in k with one positive local minimum at $k = 0$. It can be seen that the leading coefficient $a(m)$ is negative for all $m \in \Omega$. It follows that $\alpha_m(k)$ has limit to $-\infty$ when k goes to $\pm\infty$. Thus, its number of real roots is even and at least two. Suppose $\alpha_m(k)$ has at least 4 real roots, then the number of local extrema of $\alpha_m(k)$ should be at least seven, but $\alpha_m'(k)$ has degree 5, and therefore it has at most 5 roots. Therefore $\alpha_m(k)$ only has 2 real roots (one positive and one negative) and a (local) minimum at $k = 0$.

We want to see now that $\alpha_m(k)$ is positive for $k \in I$ as long as $m \in \Omega$. Note that for $m = 1$, $\alpha_1(k) = -1728k^6 + 1728k^4 + 720k^2 + 64$ is zero if and only if $k = \pm\frac{2\sqrt{3}}{3} \approx \pm 1.154 \notin I$. On the other hand, the roots of $\alpha_\omega(k)$ are ± 1. In the following claim we show that, for any $m \in [1,\omega]$, the positive root of $\alpha_m(k)$ is in the interval $\left[1, \frac{2\sqrt{3}}{3}\right]$. By symmetry, the negative root of $\alpha_m(k)$ is in $\left[-\frac{2\sqrt{3}}{3}, -1\right]$ and therefore $\alpha_m(k)$ remains positive in I, see Figure 2.5.

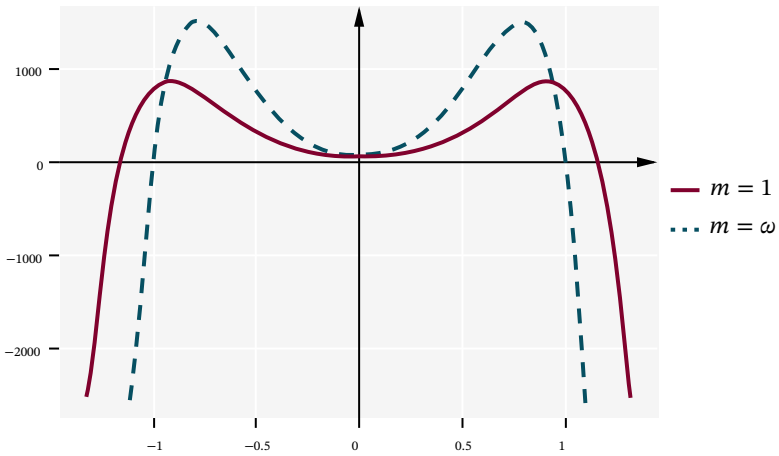


FIGURE 2.5: The solid line represents $\alpha_1(k)$, and the dashed $\alpha_\omega(k)$.

Claim. Given $m \in [1,\omega]$, there is only one positive solution $k = \mathbf{k}(m)$ of the equation $\alpha(k,m) = 0$. Moreover, the function $m \to \mathbf{k}(m)$ is continuous and strictly decreasing.

Proof of claim. As observed above, $\alpha_m(k)$ has exactly one real positive root for any $m > 1$. Note that $\mathbf{k}(m)$ is continuous because the roots of $\alpha_m(k)$ are continuous functions on the coefficients $a(m)$, $b(m)$, $c(m)$ and $d(m)$ and those coefficients are continuous on m. We have seen above that $\mathbf{k}(1) = \frac{2\sqrt{3}}{3} > 1 = \mathbf{k}(\omega)$. If $\mathbf{k}$ was not strictly decreasing, then $\mathbf{k}$ would not be injective: there would exist some $m', m'' \in [1, \omega]$ such that $\mathbf{k}(m') = \mathbf{k}(m'')$ and $\alpha(\mathbf{k}(m'), m') = \alpha(\mathbf{k}(m'), m'') = 0$. In order to reach a contradiction, we show that for any value of k, $\alpha(k, m)$ only vanishes for a unique real value of m. To this end, consider $\alpha(k, m)$ as a function of m,

$$\alpha_k(m) = \underbrace{\left(-729k^6\right)}_{a(k)} m^3 + \underbrace{243\left(-3k^6 + 4k^4 + 3k^2\right)}_{b(k)} m^2 + \underbrace{27\left(-9k^6 + 24k^4 + 2k^2\right)}_{c(k)} m$$
$$\underbrace{-27k^6 + 108k^4 - 63k^2 + 64}_{d(k)}.$$

This exhibits $\alpha_k(m)$ as a degree 3 polynomial in m and it has a unique real root since it has negative discriminant for any $k \neq 0$:

$$\begin{aligned} D(m) &= 18a(k)b(k)c(k)d(k) - 4b(k)^3d(k) + b(k)^2c(k)^2 - 4a(k)c(k)^3 - 27a(k)^2d(k)^2 \\ &= -99179645184(k^6 + 3k^8). \end{aligned}$$

Hence, we conclude that the value $\mathbf{k}(m)$ is well defined and that $m \to \mathbf{k}(m)$ is a strictly decreasing function on $[1, \omega]$.

It remains to see that $\gamma(k, m)$ does not vanish for any $(k, m) \in I \times \Omega$. Observe that $\gamma(k, m) = 0$ if and only if

$$9k\,(3m + 1) = -\sqrt{\alpha(k, m)}. \tag{2.6}$$

By squaring both members, we derive that $\alpha(k, m) - \left(9k\,(3m + 1)\right)^2 = 0$. The left member of this expression is equal to $-\left(9k^2m + 3k^2 - 4\right)^3$, which vanishes if and only if $k = \pm\frac{2}{\sqrt{9m+3}}$. Only the negative solution of k satisfies equation (2.6). Note that $k = -\frac{2}{\sqrt{9m+3}}$ is always negative and it will be smaller than $-1/3$ if and only if $m < 11/3$.

Therefore, for all $(k, m) \in I \times \Omega$, $\gamma(k, m)$ does not vanish. □

Theorem 2.4.3. *If $(k, m) \in I \times \Omega$, then $\mathbf{x}^*$ is a local minimum of the optimization problem 2.4.1 for any $T \in \mathcal{T}_4$.*

Proof. In order to prove that $\mathbf{x}^*$ is a local minimum we first show that $\mathbf{x}^*$ satisfies the Karush-Kuhn-Tucker (KKT) conditions defined in Section 2.3 for some KKT multipliers $\mu_{1,i}, \mu_{2,i}, i = 1, \dots, 5$.

1st case. Suppose $\tilde{x}(k,m) < 1$. Then we observe that $\partial_{x_1} f_{12|34}(\mathbf{x}^*) = \partial_{x_3} f_{12|34}(\mathbf{x}^*) = 0$. Moreover we have $g_{1,i}(\mathbf{x}^*) = 0$ for $i = 2, 4, 5$, $g_{1,i}(\mathbf{x}^*) \neq 0$ for $i = 1, 3$ and $g_{2,i}(\mathbf{x}^*) \neq 0$ for $i = 2, 4, 5$ and for $i = 1, 3$, $g_{2,i}(\mathbf{x}^*)$ may be zero for some values of k, m.

Suppose $g_{2,i}(\mathbf{x}^*) = 0$ for $i = 1, 3$, therefore, by (iii) of the KKT conditions, we need to take

$$\begin{aligned} \mu_{1,i} &= 0, \text{ for } i = 1, 3 \\ \mu_{2,i} &= 0, \text{ for } i = 2, 4, 5. \end{aligned}$$

Moreover, $\nabla g_{1,i}(\mathbf{x}) = (0, \dots, \overset{i}{\overbrace{1}}, \dots, 0)^t$ and $\nabla g_{2,i}(\mathbf{x}) = (0, \dots, \overset{i}{\overbrace{-1}}, \dots, 0)^t$ for all i and for every $\mathbf{x}$. Therefore condition (i),

$$\begin{aligned} -\nabla f_{12|34}(\mathbf{x}^*) = \mu_{1,2}\nabla g_{1,2}(\mathbf{x}^*) + \mu_{1,4}\nabla g_{1,4}(\mathbf{x}^*) + \mu_{1,5}\nabla g_{1,5}(\mathbf{x}^*) \\ +\mu_{2,1}\nabla g_{2,1}(\mathbf{x}^*) + \mu_{2,3}\nabla g_{2,3}(\mathbf{x}^*), \end{aligned}$$

is equivalent to

$$\left(0, \partial_{x_2} f_{12|34}(\mathbf{x}^*), 0, \partial_{x_4} f_{12|34}(\mathbf{x}^*), \partial_{x_5} f_{12|34}(\mathbf{x}^*)\right)^t = -(-\mu_{2,1}, \mu_{1,2}, -\mu_{2,3}, \mu_{1,4}, \mu_{1,5})^t,$$

which implies that necessarily

$$\begin{aligned} &\mu_{1,2} = -\partial_{x_2} f_{12|34}(\mathbf{x}^*); \quad \mu_{1,4} = -\partial_{x_4} f_{12|34}(\mathbf{x}^*); \quad \mu_{1,5} = -\partial_{x_5} f_{12|34}(\mathbf{x}^*); \\ &\mu_{2,1} = 0; \qquad\qquad\qquad \mu_{2,3} = 0. \end{aligned}$$

Observe that for the cases where $g_{2,1}(\mathbf{x}^*), g_{2,3}(\mathbf{x}^*) \neq 0$ we also have $\mu_{2,1} = 0$ and $\mu_{2,3} = 0$. Then, because of condition (iii) in KKT, it is enough to show that these three partial derivatives are negative. This is proven in the forthcoming Lemma 2.4.5.

As a consequence of these partial derivatives being negative, the entries of any directional derivative $\partial_{\mathbf{v}} f_{12|34}(\mathbf{x}^*)$ are less than or equal to zero for any vector $\mathbf{v}$. Moreover $\partial_{\mathbf{v}} f_{12|34}(\mathbf{x}^*)$ is the zero vector if and only if $\mathbf{v}$ belongs to the x_1x_3-plane. As according to Lemma 2.4.7, $\mathbf{x}^*$ is a local minimum if we fix $x_2 = x_4 = x_5 = 1$, we can conclude that $\mathbf{x}^*$ is a local minimum of $f_{12|34}$ on the region $\mathcal{D} = [-1/3, 1]^5$.

2nd case. Suppose $\tilde{x}(k,m) \geq 1$. By the KKT conditions and the same reasoning as before we need to prove that $\partial_{x_i} f_{12|34}(\mathbf{x}^*)$ is negative for every i, since no partial derivative of $f_{12|34}$ vanishes on $\mathbf{x}^*$. This is also proven in Lemma 2.4.5. Moreover, as a consequence of Lemma 2.4.7 the partial derivatives with respect to x_1 and x_3 are also negatives. Therefore $\mathbf{x}^*$ is a local optimum.

The proof for topologies $T_{13|24}$ and $T_{14|23}$ follows directly from the previous results since the functions $f_{13|24}$ and $f_{14|23}$ satisfy $\partial_{x_i} f_{13|24}(\mathbf{x}^*) = \partial_{x_i} f_{14|23}(\mathbf{x}^*) = \partial_{x_i} f_{12|34}(\mathbf{x}^*)$ for $i \neq 5$ and $\partial_{x_5} f_{13|24}(\mathbf{x}^*)$ and $\partial_{x_5} f_{14|23}(\mathbf{x}^*)$ are also negative by Lemma 2.4.6. □

In the remaining part of this section we state the technical results needed to complete the proof of Theorem 2.4.3. The proofs of these results rely on the tools and techniques of *Elimination Theory* introduced in Section 1.2.3.

In these proofs we distinguish 2 cases, depending on whether $\tilde{x}(k,m) < 1$ or $\tilde{x}(k,m) \geq 1$. For this reason, in the following lemma we start by studying for which parameters k and m one has $\tilde{x}(k,m) \geq 1$.

Lemma 2.4.4. *Let $(k,m) \in I \times \Omega$, $k \neq 0, \frac{-1}{3}$. It holds that $\tilde{x}(k,m) = 1$ if and only if m is equal to*

$$\mathrm{m}(k) := \frac{-3k^2 - k + 16}{3k(3k+1)}.$$

Moreover, $\tilde{x}(k,m) > 1$ if and only if $m > \mathrm{m}(k)$; in this case k is strictly positive.

In Figure 2.6 we represent the values $(k,m) \in I \times \Omega$ for which $\tilde{x}(k,m) \geq 1$.

Proof. Consider new variables x, g and a that will allow us to make explicit the algebraic relations of $\tilde{x}(k,m)$, $\gamma(k,m)$ and $\alpha(k,m)$. Then, for $(k,m) \in I \times \Omega$, the equality $\tilde{x}(k,m) = 1$ is satisfied if and only if (k,m) is a solution to the system of equations:

$$\begin{cases} p(x) := x - 1 = 0, \\ p_{\tilde{x}}(x,g,k,m) := 36xg - 36g^2 - 9k^2m - 3k^2 + 4 = 0, \\ p_{\gamma}(g,a,k,m) := 216g^3 - 9k\,(3m+1) - a = 0, \\ p_{\alpha}(a,k,m) := a^2 - \alpha(k,m) = 0. \end{cases} \tag{2.7}$$

Polynomials $p_{\tilde{x}}$, p_{γ} and p_{α} stand for the relations introduced in Proposition 2.4.2. Define the ideal $\mathcal{I} := \big(p(x), p_{\tilde{x}}(x,g,k,m), p_{\gamma}(x,g,a,k,m), p_{\alpha}(a,k,m)\big)$ in the polynomial ring $\mathbb{C}[x,g,a,k,m]$ and compute the elimination ideal $\mathcal{I} \cap \mathbb{C}[k,m]$. According to Theorem 1.2.31, the variety $\mathcal{V}(\mathcal{I} \cap \mathbb{C}[k,m])$ is the smallest algebraic variety containing

the possible values (k,m) that correspond to points in $\mathcal{V}(\mathcal{J})$. However this inclusion is strict and there are points $(k,m) \in \mathcal{V}(\mathcal{J} \cap \mathbb{C}[k,m])$ that do not extend to solutions of (2.7).

In this case, the ideal $\mathcal{J} \cap \mathbb{C}[k,m]$ is generated by the polynomial

$$\left(9k^2m + 3k^2 - 4\right)^3 \left(9k^2m + 3k^2 + 3km + k - 16\right). \tag{2.8}$$

This polynomial vanishes if and only if one of the factors does. The first factor $9k^2m + 3k^2 - 4$ (as a polynomial in m) has a root at $m = \frac{4-3k^2}{9k^2}$ and substituting it at (2.7) we get that either

$$\begin{cases} g = 0,\ a = -\frac{12}{k} \text{ and } k \neq 0, \text{ or} \\ g = 1,\ a = 108 \text{ and } k = \frac{1}{9}. \end{cases} \tag{2.9}$$

None of these two solutions are satisfied for $(k,m) \in I \times \Omega$. Indeed, by Lemma 2.4.2, $\gamma(k,m)$ is different from zero for $(k,m) \in I \times \Omega$, so g can not be equal to zero. The second solution in (2.9) implies $m = \frac{107}{3}$, which is not in Ω.

The second factor of the polynomial in (2.8) vanishes at the points $(k, \mathrm{m}(k))$. By Proposition 2.4.2, $\tilde{x}$ is a continuous real function on (k,m) in $I \times \Omega$. Then, in order to verify when $\tilde{x}(k,m)$ is greater than 1 it is enough to evaluate it at some point $(k,m) \in I \times \Omega$ such that $m > \mathrm{m}(k)$, and at some point $(k,m) \in I \times \Omega$ such that $m < \mathrm{m}(k)$. For example, $\tilde{x}(0, 3/2) = 0 < 1$ and $\tilde{x}(1, 3/2) \approx 1.194 > 1$. Therefore, $\tilde{x} > 1$ if and only if $m > m(k)$. Straightforward computations show that if $(k,m) \in I \times \Omega$ satisfies $\tilde{x}(k,m) \geq 1$, then $k > 0$ (see Figure 2.6). In particular, this implies that the denominator of $m(k)$ does not vanish. □

The aim of the following lemma is to prove that $\partial_{x_i} f_{12|34}(\mathbf{x}^*) < 0$ for $i = 2, 4, 5$. For each partial the idea of the proof is the same: we consider an ideal $\mathcal{J}$ for which the contraction in $\mathbb{C}[k,m]$ is the set of points (k,m) such that $\partial_{x_i} f_{12|34}(\mathbf{x}^*) = 0$ and from the study of all these points we deduce our claim.

Lemma 2.4.5. *For all $(k,m) \in I \times \Omega$ we have that*

(a) $\partial_{x_2} f_{12|34}(\mathbf{x}^*) < 0$,

(b) $\partial_{x_4} f_{12|34}(\mathbf{x}^*) < 0$,

(c) $\partial_{x_5} f_{12|34}(\mathbf{x}^*) < 0$.

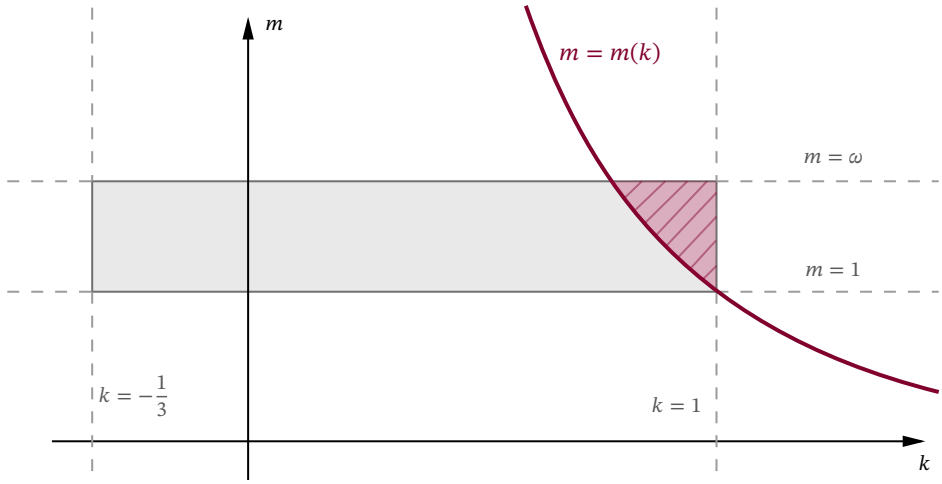


FIGURE 2.6: The red curve represents the function m(k) and the grey rectangle is $I \times \Omega$. The stripped region correspond to the values $(k, m) \in I \times \Omega$ such that $\tilde{x}(k, m) \geq 1$ (see Lemma 2.4.4).

Proof. Given $(k, m) \in I \times \Omega$, write $\tilde{x}$ for $\tilde{x}(k, m)$.

(a) The proof falls naturally into two cases.
1st case. Suppose $\tilde{x} < 1$. By definition, $\mathbf{x}^* = (\tilde{x}, 1, \tilde{x}, 1, 1)$ in this case. Therefore, $\partial_{x_2} f_{12|34}((x, 1, x, 1, 1))$ is given by the polynomial:

$$p(x, k, m) := 54x^4 - 18\left(2k^2m + k^2 - 2\right)x^2 - 6\left(5km + k\right)x - 6m + 6.$$

In order to prove that this function is negative we prove that it never vanishes on $I \times \Omega$ and is negative for a particular value in that region. $\partial_{x_2} f_{12|34}(\mathbf{x}^*)$ is zero if and only if the following polynomials vanish:

$$p(x, k, m),\ p_{\tilde{x}}(x, g, k, m),\ p_\gamma(g, a, k, m),\ \text{and } p_\alpha(a, k, m), \tag{2.10}$$

where $p_{\tilde{x}}(x, g, k, m)$, $p_\gamma(g, a, k, m)$ and $p_\alpha(a, k, m)$ are defined as in (2.7).

We consider the ideal $\mathcal{J} = \left(p(k, m, x), p_{\tilde{x}}(k, m, x, g), p_\gamma(k, m, x, g, a), p_\alpha(k, m, a)\right)$ and we compute the elimination ideal $\mathcal{J} \cap \mathbb{C}[k, m]$ which turns out to be generated by exactly one polynomial:

$$(m - 1)\left(3k^2 + 1\right)\left(9k^2m + 3k^2 - 4\right)^3 h(k, m) \tag{2.11}$$

where

$$h(k,m) = 81k^6m^3 - 27k^6m^2 - 45k^6m - 9k^6 + 39k^4m^3 + 547k^4m^2 + 469k^4m + 97k^4 - 1312k^2m^2 - 1120k^2m - 256k^2 - 768m^2.$$

The polynomial in (2.11) is zero if and only if at least one of its factors vanishes. The first factor is zero when $m = 1$, but $1 \notin \Omega$. The second one has no real solutions in k. Note that $9k^2m + 3k^2 - 4$ is zero when $k = \pm\frac{2}{\sqrt{9m+3}}$. However, the negative solution does not belong to I if $m \in \Omega$ (see the proof of Lemma 2.4.2) and the positive one does not generate a solution to (2.10) since $\tilde{x}\left(\frac{2}{\sqrt{9m+3}}, m\right) = \sqrt[6]{\frac{3m+1}{108}} \neq 0$ for $m \in \Omega$. The case of $h(k,m)$ is not that simple. Consider h as a polynomial in m:

$$h_k(m) = \underbrace{\left(81k^6 + 39k^4\right)}_{a(k)} m^3 + \underbrace{\left(-27k^6 + 547k^4 - 1312k^2 - 768\right)}_{b(k)} m^2 + \underbrace{\left(-45k^6 + 469k^4 - 1120k^2\right)}_{c(k)} m + \underbrace{\left(-9k^6 + 97k^4 - 256k^2\right)}_{d(k)}.$$

The discriminant of $h_k(m)$ is

$$D(k) = -49152k^2(\sqrt{6} - k)(\sqrt{6} + k)(384 - 106k^2 + 39k^4)(64 + 115k^2 - 38k^4 + 3k^6)^2.$$

The discriminant $D(k)$ has three real roots at $k = 0$ and $k = \pm\sqrt{6}$ with $\sqrt{6} \sim 2.449$. Since $D(-1) = D(1) < 0$ we conclude $D(k) \leq 0\ \forall k \in I$ and hence $h_k(m)$ only has one real root in this interval. Since the leading coefficient of h_k is positive and $h_k(2) = 441k^6 + 3535k^4 - 7744k^2 - 3072 < 0$ we conclude that the root of $h_k(m)$ is greater than 2 and therefore does not belong to Ω.

Consequently there are no points in $\mathcal{V}(\mathcal{I} \cap \mathbb{C}[k,m])$ in the region $I \times \Omega$. Since $\partial_{x_2} f_{12|34}(\mathbf{x}^*)$ is continuous and well defined in $I \times \Omega$ we conclude that $\partial_{x_2} f_{12|34}(\mathbf{x}^*)$ has the same sign in all the domain. Evaluating at any point $(k,m) \in I \times \Omega$ we deduce that $\partial_{x_2} f_{12|34}(\mathbf{x}^*)$ is negative on this region.

2nd case. Suppose that $\tilde{x} \geq 1$. We already know that in this case, $m \geq \mathrm{m}(k)$, which implies that $k > 0$ (see Lemma 2.4.4). On the other hand, we have $\partial_{x_2} f_{12|34}(\mathbf{1}) = -18k^2 - 6(6k^2 + 5k + 1)m - 6k + 96$ is negative if and only if $m > \frac{-3k^2-k+16}{6k^2+5k+1}$. Now, it is straightforward to check that for positive k, $\mathrm{m}(k) > \frac{-3k^2-k+16}{6k^2+5k+1}$.

(b) Computing the partial derivative with respect to x_4 and substituting at $\mathbf{x}^*$ we get

$$\partial_{x_4} f_{12|34}(\mathbf{x}^*) = \partial_{x_2} f_{12|34}(\mathbf{x}^*),$$

which follows from the symmetry on $f_{12|34}$ and on $\mathbf{x}^*$. Then, (b) is a consequence of (a).

(c) We also split the proof of $\partial_{x_5} f_{12|34}(\mathbf{x}^*) < 0$ into two cases

1st case. Suppose $\tilde{x} < 1$:

$$\partial_{x_5} f_{12|34}(\mathbf{x}^*) = 54\tilde{x}^4 - 18(3k^2m - 2)\tilde{x}^2 - 36km\tilde{x} - 6m + 6.$$

In this case consider the ideal $\mathcal{J} = \big(p(x,k,m), p_{\tilde{x}}(x,g,k,m), p_\gamma(x,g,a,k,m), p_\alpha(a,k,m)\big)$ where $p(x,k,m) = 54x^4 - 18(3k^2m-2)x^2 - 36kmx - 6m + 6$. The ideal $\mathcal{J} \cap \mathbb{C}[k,m]$ is generated by the polynomial

$$(m-1)\left(3k^2+1\right)\left(9k^2m + 3k^2 - 4\right)^3 h(k,m),$$

where $h(k,m) = 81k^4m^3 - (27k^4 + 288k^2 + 256)m^2 - (45k^4 + 96k^2)m - 9k^4$. We only need to study the intersection of $h(k,m)$ with $I \times \Omega$ since the other factors have already been studied in the proof of (a). Taking $h(k,m)$ as a function of m we compute its discriminant,

$$D(k) = -442368k^6(2 + 3k^2)(128 + 18k^2 + 27k^4)$$

which has only one real root at $k = 0$. Substituting at $k = \pm 1$ we get $D(-1) = D(1) = -382648320 < 0$. Therefore $D(k) \leq 0\ \forall k \in I$ and $h_k(m)$ has exactly one real root. If $k \in I$ this root is not in Ω since $h(k,1) = -384k^2 - 256 < 0\ \forall k$, and $h(k,2) = 441k^4 - 1344k^2 - 1024 < 0\ \forall k \in I$. Then, is valid to conclude $\partial_{x_5} f_{12|34}(\mathbf{x}^*)$ is negative in our domain.

2nd case. Suppose $\tilde{x} \geq 1$: The function $\partial_{x_5} f_{12|34}(\mathbf{1}) = -6(9k^2 + 6k + 1)m + 96$ is negative if and only if $m > \frac{16}{9k^2+6k+1}$. The value $\mathrm{m}(k)$ defined in Lemma 2.4.4 is greater than $\frac{16}{9k^2+6k+1}$ for all $k \in [0,1]$. Since $k > 0$ when $\tilde{x}(k,m) > 1$, $\partial_{x_2} f_{12|34}(\mathbf{1})$ is negative for all $k \in I, m \in \Omega$ such that $m > \mathrm{m}(k)$. □

Lemma 2.4.6. *For all $(k,m) \in I \times \Omega$ we have that*

(a) $\partial_{x_5} f_{13|24}(\mathbf{x}^*) \leq 0$,

(b) $\partial_{x_5} f_{14|23}(\mathbf{x}^*) \leq 0$.

Proof. (a) We split the proof into two cases as above.

1st case. Assume $\tilde{x} < 1$, then:

$$\partial_{x_5} f_{13|24}(\mathbf{x}^*) = 48\tilde{x}^4 - 12(3k^2m + k^2 - 4)\tilde{x}^2 - 12(3km + k)\tilde{x}.$$

Write $p(x, k, m)$ for this polynomial and

$$\mathcal{J} = \left(p(x, k, m), p_{\tilde{x}}(x, g, k, m), p_\gamma(x, g, a, k, m), p_\alpha(a, k, m)\right).$$

In this case the elimination ideal $\mathcal{J} \cap \mathbb{C}[k, m]$ is generated by the polynomial

$$k^4(m-1)(3m+1)^3(9k^2m + 3k^2 - 4)^3$$

which vanishes if and only if $m = 1$, $m = -1/3$, $k = 0$ or $m = \frac{4-3k^2}{9k^2}$. The two first possible values of m do not belong to Ω. If $m = \frac{4-3k^2}{9k^2}$, then $\partial_{x_5} f_{13|24}(\mathbf{x}^*)$ vanishes if and only if $k = 1/\sqrt{3}$, but then $m = 1$, which is not in Ω. It only remains to study the case $k = 0$. Having $k = 0$ implies that $\tilde{x} = 0$ and therefore $\partial_{x_5} f_{13|24}(\mathbf{x}^*) = 0$. However, if we evaluate $\partial_{x_5} f_{13|24}(\mathbf{x}^*)$ at $k = 1$ and $k = -1$, we check that it takes negative values. Therefore, since the derivative $\partial_{x_5} f_{13|24}(\mathbf{x}^*)$ only vanishes at $k = 0$ (for $(k, m) \in I \times \Omega$) we deduce our claim.

2nd case. Suppose $\tilde{x} \geq 1$: The value of $\partial_{x_5} f_{13|24}(\mathbf{1}) = -6(9k^2 + 6k + 1)m + 96$ is negative if and only if $m > \frac{16}{9k^2-6k+1}$. Since the value $\mathrm{m}(k)$ obtained in Lemma 2.4.4 is greater than $\frac{16}{9k^2-6k+1}$ for all $k \in [0, 1]$, the claim follows.

(b) Consider also two cases.

1st case. Assume $\tilde{x} < 1$, then

$$\partial_{x_5} f_{14|23}(\mathbf{x}^*) = 54\tilde{x}^4 - 6(7k^2m + 2k^2 - 6)\tilde{x}^2 - 12(2km + k)\tilde{x} - 6m + 6;$$

write $p(x, k, m)$ for this polynomial. If

$$\mathcal{J} := \left(p(x, k, m), p_{\tilde{x}}(x, g, k, m), p_\gamma(x, g, a, k, m), p_\alpha(a, k, m)\right),$$

then the elimination ideal $\mathcal{J} \cap \mathbb{C}[k, m]$ is generated by the polynomial

$$(m-1)(9k^2m + 3k^2 - 4)^3 h(k, m) \tag{2.12}$$

where $h(k,m) = a(m)k^8 + b(m)k^6 + c(m)k^4 + d(m)k^2 + e(m)$ and

$$\begin{aligned} a(m) &= 36m^4 - 129m^3 + 19m^2 + 61m + 13, \\ b(m) &= -942m^3 + 2362m^2 + 1750m + 286, \\ c(m) &= -2097m^3 + 7003m^2 + 3853m + 457, \\ d(m) &= 672m^2 + 8928m + 3072, \\ e(m) &= 2304m^2 \end{aligned}$$

The polynomial in (2.12) vanishes if $m = 1 \notin \Omega$, $k = \pm\frac{2}{\sqrt{9m+3}}$ or $h(k,m)$ is zero. However, recall that $k = -\frac{2}{\sqrt{9m+3}}$ does not belong to I if $m \in \Omega$ (see the proof of Lemma 2.4.2) and evaluating $\partial_{x_5} f_{14|23}(\mathbf{x}^*)$ at $k = \frac{2}{\sqrt{9m+3}}$ one can check that it vanishes if and only if $m = 1$ which is not in Ω.

It remains to check whether $h(k,m)$ vanishes for any values $(k,m) \in I \times \Omega$. Straightforward computations show that the roots of the polynomials $a(m)$, $b(m)$, $c(m)$, $d(m)$ and $e(m)$ do not lie in Ω. By evaluating these polynomials at particular values of Ω, it is immediate to check that $a(m)$ is negative, while the other polynomials are positive. Thus, by the Descartes rule, $h_m(k)$ (i.e. $h(k,m)$ considered as a function of k) has only one positive real root. Since it is an even plynomial on k is has also one real negative root. We claim that the positive root of $h_m(k)$ is greater than 1 for any $m \in \Omega$: observe that $h_m(0) = 2304m^2$ is always positive. Moreover, it is easy to check that the polynomial $h_m(1) = 36m^4 - 3168m^3 + 12360m^2 + 14592m + 3828$ is always positive for any $m \in \Omega$. Then, since $h_m(k)$ has only one positive root and $h_m(0)$, $h_m(1) > 0$, the roots of $h_m(k)$ do not lie in I for $m \in \Omega$. Evaluating $\partial_{x_5} f_{13|24}(\mathbf{x}^*)$ at any point in $I \in \Omega$, we check that it takes a negative value.

2nd case. Suppose $\tilde{x} \geq 1$: The value of $\partial_{x_5} f_{14|23}(\mathbf{1}) = -6(7k^2+4k+1)m - 12(k^2+k-8)$ is negative if and only if $m > -\frac{2(k^2+k-8)}{7k^2+4k+1}$. Moreover, $-\frac{2(k^2+k-8)}{7k^2+4k+1} < \mathrm{m}(k)$ (see Lemma 2.4.4 for a definition of $m(k)$) for all $k \in [0,1]$. Then, the statement follows.

□

Lemma 2.4.7. *For any quartet tree topology $T \in \mathcal{T}_4$ consider the function $g : I \times I \longrightarrow \mathbb{R}$ defined as $g(x,y) = f_T(x,1,y,1,1)$. Then, the point*

$$\mathbf{u} := \begin{cases} (\tilde{x}(k,m), \tilde{x}(k,m)) & \text{if } \tilde{x}(k,m) < 1; \\ (1,1) & \text{otherwise;} \end{cases}$$

is a local minimum of g. In particular, $\partial_{x_1} f_T(\mathbf{x}^*)$ and $\partial_{x_3} f_T(\mathbf{x}^*)$ are less than or equal to zero for any T.

Proof. Straightforward computations show that $f_{12|34}(x,1,y,1,1) = f_{13|24}(x,1,y,1,1) = f_{14|23}(x,1,y,1,1) = g(x,y)$. Therefore the following proof is valid for any (trivalent) tree topology with 4 leaves. In order to prove that $\mathbf{u}$ is a local minimum of $g(x,y)$ we consider two cases. We first assume that $\tilde{x} < 1$ and we will prove that $\mathbf{u}$ is a *local* minimum of g. The second case is when $\tilde{x} \geq 1$ so that $\mathbf{u}$ is on the boundary of $I \times I$; by the KKT conditions we must prove that $\nabla g(1,1)$ is negative. We write $\tilde{x}$ for $\tilde{x}(k,m)$.

1st case. Assume $\tilde{x} < 1$. The first derivatives of $g(x,y)$ vanish at $\mathbf{u}$. The Hessian matrix of g evaluated at a point (x,x) is

$$\mathbf{H} = \begin{pmatrix} 72x^2+24 & -54k^2m - 18k^2 + 144x^2 \\ -54k^2m - 18k^2 + 144x^2 & 72x^2+24 \end{pmatrix}.$$

To show that $\mathbf{H}$ is a positive definite matrix, we see that all its leading principal minors are positive for all $(k,m) \in I \times \Omega$, see Theorem 1.2.6. The first one is clearly positive. To prove that the determinant of $\mathbf{H}$ is also positive we will follow the same ideas of the previous lemmas.

Consider the ideal $\mathcal{J} = \big(\det(\mathbf{H}), p_{\tilde{x}}(x,g,k,m), p_\gamma(x,g,a,k,m), p_\alpha(a,k,m)\big)$ where

$$\det(\mathbf{H}) = -324(3k^2m + k^2 - 8x^2)^2 + 576(3x^2+1)^2.$$

The elimination ideal $\mathcal{J} \cap \mathbb{C}[k,m]$ is generated by the polynomial

$$\big(9k^2m + 3k^2 - 4\big)^3 \big(27k^2m^2 - 126k^2m - 45k^2 - 64\big)\, h(k,m) \tag{2.13}$$

where

$$\begin{aligned} h(k,m) = {} & 729k^6m^3 + 729k^6m^2 + 243k^6m + 27k^6 - 972k^4m^2 - 648k^4m - 108k^4 \\ & - 729k^2m^2 - 54k^2m + 63k^2 - 64. \end{aligned}$$

We are interested in the real zeros of each factor of (2.13). As in the previous lemmas, it is straightforward to check that the points of the form $(k, \frac{-3k^2+4}{9k^2})$ that lie on the domain $I \times \Omega$ do not extend to solutions of the original ideal $\mathcal{J}$; more precisely, $\det(\mathbf{H})$ does not vanish over these points. For any value $k \in I$, the second factor in (2.13) only vanishes at $m = \frac{21k \pm 8\sqrt{9k^2+3}}{9k}$, which does not belong to Ω. Indeed, if we denote $m^+(k) = \frac{21k+8\sqrt{9k^2+3}}{9k}$

and $m^-(k) = \frac{21k-8\sqrt{9k^2+3}}{9k}$, then we want to prove that the image of these functions does not meet Ω. Note that $m^+(k)$ is a decreasing function since its derivative $\partial_k m^+(k) = \frac{-8}{3k^2\sqrt{9k^2+3}}$ is negative for all $k \neq 0$. Moreover,

$$\begin{aligned} &\lim_{k\to-\infty} m^+(k) = \frac{-1}{3}, \quad \lim_{k\to+\infty} m^+(k) = 5, \\ &\lim_{k\to0^-} m^+(k) = -\infty, \quad \lim_{k\to0^+} m^+(k) = +\infty. \end{aligned}$$

Hence, $Im(m^+(k)) \cap \Omega$ is empty. The function $m^-(k)$ is increasing since its derivative $\partial_k m^-(k) = \frac{8}{3k^2\sqrt{9k^2+3}}$ is positive for all $k \neq 0$. The limits of this function are

$$\begin{aligned} &\lim_{k\to-\infty} m^-(k) = 5, \quad \lim_{k\to+\infty} m^-(k) = \frac{-1}{3}, \\ &\lim_{k\to0^-} m^-(k) = +\infty, \quad \lim_{k\to0^+} m^-(k) = -\infty \end{aligned}$$

and therefore the image of $m^-(k)$ neither intersects with Ω.

Consider $h(k,m)$ as a function of m. As its discriminant $D(k) = -297538935552k^8 - 99179645184k^6$ is negative for all $k \neq 0$ the polynomial $h_k(m)$ has at most one real root $\forall k$. Moreover, $h_k(1) \leq 0$ and $h_k(\omega) \leq 0$ for all k and hence h_k is smaller or equal than zero for all $m \in \Omega$. Therefore it can be deduced that $\det(\mathbf{H})$ has constant sign in the region $I \times \Omega$. Substituting at a particular point on that region we check that $\det(\mathbf{H}) > 0$ for all $(k,m) \in I \times \Omega$.

2nd case. Assume $\tilde{x} \geq 1$. In this case, since we are in the boundary of the domain, we need to prove that $\nabla g(1,1) < 0$. The gradient

$$\nabla g(1,1) = (-54k^2m - 18k^2 - 18km - 6k + 96, -54k^2m - 18k^2 - 18km - 6k + 96)$$

is zero if and only if $m = \mathrm{m}(k)$. Moreover for $m \geq \mathrm{m}(k)$ or equivalently for $\tilde{x} \geq 1$ the polynomial $-54k^2m - 18k^2 - 18km - 6k + 96$ is negative.

Finally, for any T, $\partial_{x_1} f_T(\mathbf{x}^*) = \partial_x g(\mathbf{u})$ and $\partial_{x_3} f_T(\mathbf{x}^*) = \partial_y g(\mathbf{u})$. Therefore, as already shown the partials $\partial_{x_1} f_T(\mathbf{x}^*)$ and $\partial_{x_3} f_T(\mathbf{x}^*)$ are zero if $\tilde{x}(k,m) < 1$ and negative if $\tilde{x}(k,m) \geq 1$. □

The computations in this section have been done with `SageMath` The Sage Developers (2019) version 8.6 and `Macaulay2` (Grayson and Stillman, 2009) version 1.17.

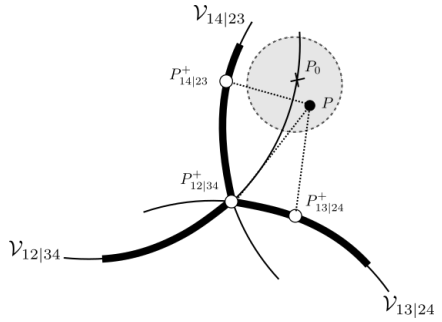


FIGURE 2.7: The point p_0 lies in the phylogenetic variety $\mathcal{V}_{T_{12|34}}$ outside the stochastic region ($m > 1$). Under the assumption of Theorem 2.4.10, as long as the point p is close to p_0, it will remain closer to $\mathcal{V}^+_{T_{13|24}}$ or $\mathcal{V}^+_{T_{14|23}}$ than to $\mathcal{V}^+_{T_{12|34}}$.

2.4.2 Global minimum

Although we are not able to prove that the local minimum presented above is indeed a global minimum, our simulation studies suggest that it is the case for $T = T_{12|34}$:

Conjecture 2.4.8. Let $T = T_{12|34}$ and $p_0 := \varphi_T(k_0, 1, k_0, 1, m_0)$. If $(k_0, m_0) \in I \times \Omega$, then

$$d(p_0, \mathcal{V}_T^+) = d\big(p_0, \varphi_T\big(\tilde{x}(k_0, m_0), 1, \tilde{x}(k_0, m_0), 1, 1\big)\big)$$

and $\varphi_T\big(\tilde{x}(k_0, m_0), 1, \tilde{x}(k_0, m_0), 1, 1\big)$ is the unique point in $\mathcal{V}_T^+$ that minimizes the distance to p_0.

Remark 2.4.9. We have tested the conjecture for 1000 pairs of parameters (k_0, m_0) randomly chosen on the region $(0, 1/4] \times (1, 3/2]$ in order to simulate points close to the LBA phenomenon. Every experiment has verified that the global minimum of the problem is unique and is the point $\mathbf{x}^*$, defined in (2.4) (which we proved to be a local minimum). The computations have been done with `Macaulay2` and a list of the tested parameters k_0 and m_0 can be found in the repository

`https://github.com/marinagarrote/StochasticPhylogeneticRegions`.

Though the conjecture is stated for $T = T_{12|34}$, note that it would also be true for any $T \in \mathcal{T}_4$ by permuting the parameters accordingly.

In the following theorem, we assume $T = T_{12|34}$ and prove that, for any point p close enough to a point $p_0 = \varphi_T(k_0, 1, k_0, 1, m_0)$ satisfying the previous conjecture, the distance from p to the stochastic phylogenetic region $\mathcal{V}^+_{T'}$, for $T' \neq T$, is upper bounded by the distance from p to $\mathcal{V}^+_T$; see Figure 2.7 for an illustration.

Theorem 2.4.10. *Let $T = T_{12|34}$ and $p_0 := \varphi_T(k_0, 1, k_0, 1, m_0) \in \mathcal{V}_T$ with $(k_0, m_0) \in I \times \Omega$. Assume that the minimum distance from p_0 to $\mathcal{V}_T^+$ is attained at a unique point p_0^+ given by $p_0^+ = \varphi_T\big(\tilde{x}(k_0, m_0), 1, \tilde{x}(k_0, m_0), 1, 1\big)$ with $\tilde{x}(k_0, m_0) \neq 0$. If p is close enough to p_0 and $T' \neq T$ is another tree in $\mathcal{T}_4$, then the closest point in $\mathcal{V}_T^+$ (that is p_T^+) belongs also to $\mathcal{V}_{T'}^+$. In particular,*

$$d(p, \mathcal{V}_T^+) \geq d(p, \mathcal{V}_{T'}^+).$$

Proof. We consider the following sets of points in the border of $\mathcal{V}_T^+$,

$$\begin{aligned} \mathcal{B}_1 &:= \varphi_T(S_{1,5}) \\ \mathcal{B}_2 &:= \varphi_T(\mathcal{D} \cap \{x_5 = -1/3\}). \end{aligned}$$

Given a point p, we define $f_{T,p}(\mathbf{x})$ as the square of the distance function from $\varphi_T(\mathbf{x})$ to p, and we consider the set

$$W_p := \left\{ \varphi_T(\mathbf{x}) \;\middle|\; \mathbf{x} \in \mathcal{D} \text{ and } \partial_{x_5} f_{T,p}(\mathbf{x}) = 0 \right\} \cup \mathcal{B}_2.$$

Define also $g(p) := d(p, W_p) - d(p, \mathcal{B}_1)$, which is continuous as a function of p.

By hypothesis, p_0^+ equals $\varphi_T(\mathbf{x}_0^*)$, where $\mathbf{x}_0^* = \big(\tilde{x}(k_0, m_0), 1, \tilde{x}(k_0, m_0), 1, 1\big)$. Since $\tilde{x}(k_0, m_0) \neq 0$, $\mathbf{x}_0^*$ is the *only* preimage in $\mathcal{D}$ of p_0^+ (see Casanellas and Fernández-Sánchez, 2008). Therefore, p_0^+ lies in $\mathcal{B}_1$ but not in W_{p_0} (see Lemma 2.4.5), so that $g(p_0) > 0$. If p is close enough to p_0, then the value $g(p)$ remains positive. This implies that the global minimum p_T^+ (which is still unique if p is close to p_0) of $f_{T,p}(\mathbf{x})$ lies in $\varphi_T(\mathcal{D})$ does not lie in W_p and therefore lies in the border $\mathcal{B}_1$. As a consequence, p_T^+ lies also in $\mathcal{V}_{T'}^+$ and $d(p, \mathcal{V}_T^+) \geq d(p, \mathcal{V}_{T'}^+)$. □

Remark 2.4.11. With the notation of the previous theorem, note that when $k_0 = 1$, this situation is a special case of the situation considered in Section 2.2; the only difference is that we are now restricting ourselves to the JC69 model instead of considering the K81. The result obtained here coincides with the case considered in Prop. 2.2.1 (c), where the closest point lies in the intersection of the varieties and $d(P, \mathcal{V}_T^+) = d(p, \mathcal{V}_{T'}^+)$.

2.5 Study on simulated data

In this section we simulate points close to a given phylogenetic variety and we compute its distance to the stochastic region of this variety and to the other phylogenetic varieties (distinguishing also their stochastic regions). The computations are performed in the

setting of long branch attraction of the previous section and also for balanced trees. We cannot do this theoretically because, even if we have found a local minimum for the long branch attraction setting (Theorem 2.4.3), we cannot guarantee that it is global and also because we do not have a formula for the distance when the input does not lie on the variety. The computations of this section are performed using Algorithm 2.2; the implementation can be found in Code A.2 and A.3 and can be downloaded from the github repository mentioned in Remark 2.4.9.

We consider the quartet $T = T_{12|34}$ displayed in Figure 2.8 evolving with JC69 matrices. As presented in the figure, we suppose the matrix K_a is attached to the edges adjacent to leaves 1 and 3 and K_b to edges ending at leaves 2 and 4. We suppose both K_a and K_b are stochastic JC69 matrices with eigenvalues different from one k_a and k_b, respectively. The matrix M is a JC69 matrix attached at the interior edge of T, with eigenvalue m that takes values in the interval $[0.94, 1.06]$ by steps of length 0.02. We take $p = \varphi_T(k_a, k_b, k_a, k_b, m)$ which are points in $\mathcal{V}_{T_{12|34}}$ that range from the stochastic region of the variety $\mathcal{V}^+_{T_{12|34}}$ (that is $m \leq 1$) to the non-stochastic part ($m > 1$). We consider two different scenarios:

- trees satisfying the LBA penomenon, that is $k_a = 0.37$ and $k_b = 0.87$,
- balanced trees with $k_a = k_b = 0.51$.

The set of parameters k_a, k_b and m have been chosen in order that $\varphi_T(k_a, k_b, k_a, k_b, m)$ is always a distribution. Then, for each set of parameters, we considered 100 data points, each corresponding to the observation of 10000 independent samples from the corresponding multinomial distribution $\varphi_T(k_a, k_b, k_a, k_b, m)$. As the varieties $\mathcal{V}_T$ all lie in a linear space of dimension 14 (see Example 1.3.20), we first project these data to this linear variety.

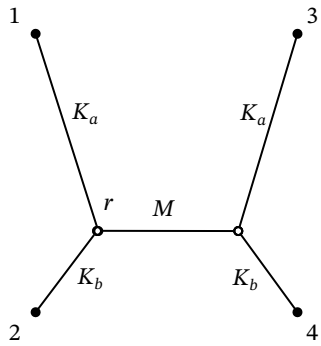


FIGURE 2.8: Quartet $T = T_{12|34}$ with distribution $p = \varphi(k_a, k_b, k_a, k_b, m)$ arising from T with the stochastic JC69 matrices K_a and K_b at the exterior edges and the JC69 matrix M (not necessarily stochastic) at the interior edge.

For each data point p generated as above and for each tree $T \in \mathcal{T}_4$, we have computed the distance from p to the stochastic region $\mathcal{V}_T^+$, $d(p, \mathcal{V}_T^+)$ using Algorithm 2.2 and the distance to the phylogenetic variety, $d(p, \mathcal{V}_T)$. These computations have been performed for the three tree topologies $T_{12|34}$, $T_{13|24}$ and $T_{14|23}$.

For each set of parameters k_a, k_b and m we have plotted the average of each of these distances computed from the 100 data points. In each graphic we have fixed k_a and k_b and let m vary in the x-axis from 0.94 to 1.06; the y-axis represents the distance. The grey background part of the plots represents the region of data points sampled from non-stochastic parameters, whereas the white part represents the stochastic region.

The plots on the top of Figure 2.9 represent trees in the long branch attraction (LBA) case (see Figure 2.7), while those on the bottom represent balanced trees; on the left we represent the distance to the phylogenetic varieties and on the right to the stochastic phylogenetic regions. Concerning the plots on the left (distance to the phylogenetic varieties), the distance to $\mathcal{V}_{T_{12|34}}$ is always smaller for balanced trees (for all values of m), but this does not hold in the LBA case (top left figure): for points close to the intersection of the varieties, that is, m close 1, the points are closer to variety corresponding to the tree $T_{13|24}$ (this is the reason why methods based solely on algebraic tools might perform incorrectly in the LBA case). In both cases (long branch attraction and balanced trees) we observe a similar behaviour on the plots on the right (distance to stochastic regions): we note that for $m \leq 1$ the distance to $\mathcal{V}_{T_{12|34}}^+$ is almost always the smallest (except for some points with m very close to 1 in the top figure) and when $m > 1$ the distance to $\mathcal{V}_{T_{12|34}}^+$ becomes greater than the distance to the other stochastic regions. This illustrates the inequality of Theorem 2.4.10.

The different performance on the two plots of the distances to $\mathcal{V}_{T_{13|24}}^+$ and $\mathcal{V}_{T_{14|23}}^+$ are due to the shapes of the trees that we are considering. When the tree is balanced we see that the distances to $\mathcal{V}_{T_{13|24}}^+$ and $\mathcal{V}_{T_{14|23}}^+$ are almost equal.

Every simulation performed has showed us that, when $m > 1$, the closest point to p in $\mathcal{V}_{T_{12|34}}^+$ (that is, $p_{12|34}^+$) belongs to the intersection of the varieties, i.e. $p_{12|34}^+ \in \mathcal{V}_{T_{12|34}}^+ \cap \mathcal{V}_{T_{13|24}}^+ \cap \mathcal{V}_{T_{14|23}}^+$. However, this is not true when we compute the closest point to $\mathcal{V}_{T'}^+$ for $T' \neq T_{12|34}$. In the case of long branch attraction the closest point $p_{14|23}^+ \in \mathcal{V}_{T_{14|23}}^+$ to p was always the image of parameters at the interior of $\mathcal{D}$ by $\varphi_{14|23}$ whether for $T = T_{13|24}$, the parameters giving rise to $p_{13|24}^+$ were in the interior of $\mathcal{D}$ approximately half of the time.

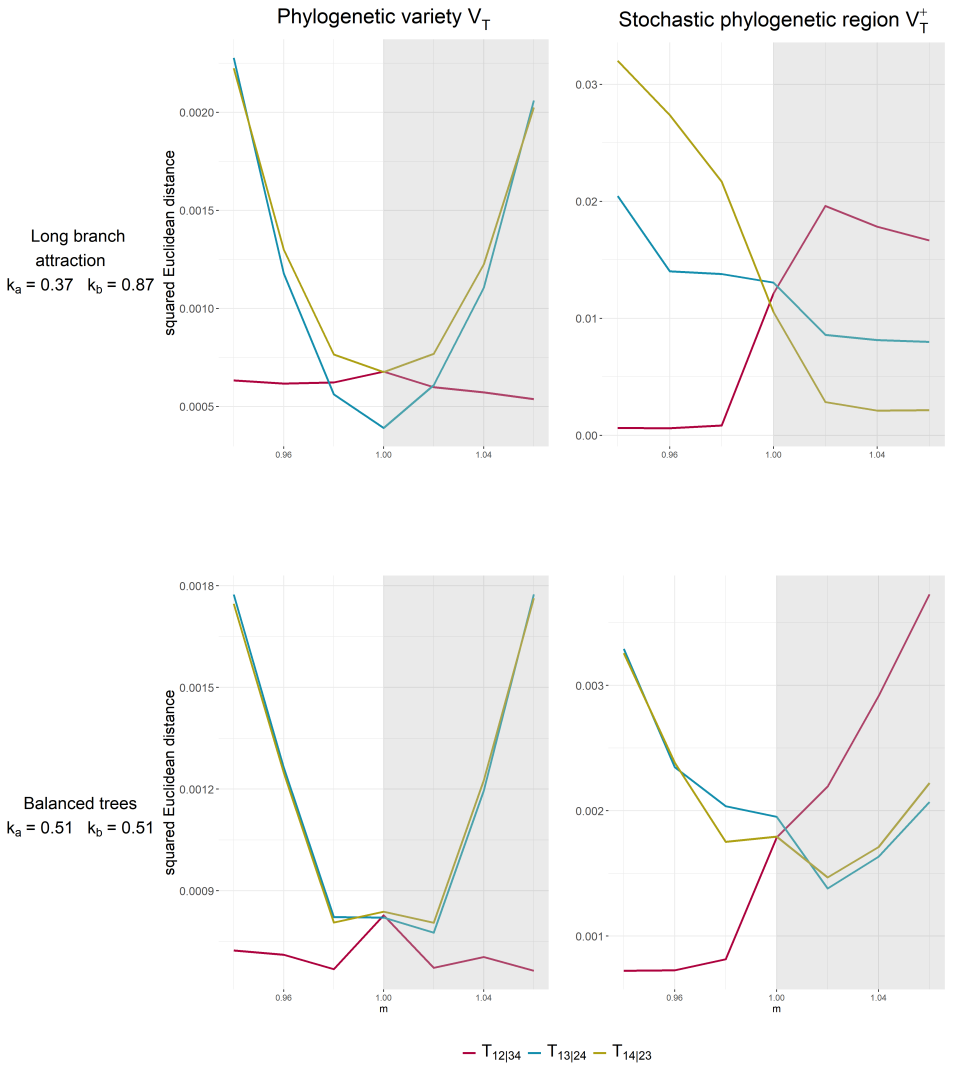


FIGURE 2.9: These four plots represent the distance of sampled points to the phylogenetic varieties (on the left) and to their stochastic region (on the right). In each plot, the horizontal axis represents the eigenvalue m of the matrix M in the tree of Figure 2.8. The two plots on top correspond to the long branch attraction situation, while the two plots on bottom correspond to balanced trees. The grey background part indicates the values of m for which M is not a stochastic matrix.

These simulations verify that, if $p \in \mathbb{R}^{4^4}$ is a distribution satisfying $d\left(p, \mathcal{V}_{T_{12|34}}\right) < \min\left\{d\left(p, \mathcal{V}_{T_{13|24}}\right), d\left(p, \mathcal{V}_{T_{14|23}}\right)\right\}$, it is possible that

$$d\left(p, \mathcal{V}^+_{T_{12|34}}\right) > \min\left\{d\left(p, \mathcal{V}^+_{T_{13|24}}\right), d\left(p, \mathcal{V}^+_{T_{14|23}}\right)\right\}.$$

This provides an affirmative answer to the Question 2.1. This suggests that considering the stochastic part of phylogenetic varieties and the resulting semi-algebraic constraints needed to describe them may be an interesting strategy for phylogenetic reconstruction in the long branch attraction setting, and also for balanced trees. However, as it has become evident throughout this chapter, dealing with both algebraic and semi-algebraic conditions is not an easy task, and more work is needed in order to design practical methods for phylogenetic inference under more general evolutionary models than the models used here. This shall be addressed in the following chapter.

Computations

The computations of this section were performed on a machine with 10 Dual Core Intel(R) Xeon(R) Silver 64 Processor 4114 (2.20 GHz, 13.75M Cache) equipped with 256 GB RAM running Ubuntu 18.04.2. We have used `Macaulay2` version 1.3 and `SageMath` version 8.6.

3

New Phylogenetic reconstruction methods

In this chapter we study how to combine the algebraic and semi-algebraic descriptions of the general Markov model (see Theorems 1.3.24 and 1.3.26) to design a method that allows the reconstruction of the tree topology taking sequence alignments as input.

To this end, we first need to guarantee that the rank does not increase when projecting a matrix into the space of positive definite matrices. Section 3.1 is devoted to prove this technical result. In Section 3.2 we describe and analyze the *leaf-transformations* on some particular tensors, which will be crucial in the development of our methods. The first reconstruction method of the chapter is called `SAQ` (named for *Semi-Algebraic Quartet reconstruction method*) and is introduced and described in Section 3.3. In particular, in Theorem 3.3.1 we prove its statistical consistency. In Section 3.4 we describe a criterion based on the paralinear distance that will be fundamental in the analysis of the data point obtained from the sequence alignment. This criterion is applied in Section 3.5 in the design of our second reconstruction method, `ASAQ` (for *Algebraic and Semi-Algebraic Quartet reconstruction method*), which takes advantage of `SAQ` and `Erik+2` and becomes a powerful phylogenetic reconstruction method.

3.1 The inertia of the PSD approximation

In this section, we focus on the study of the rank and inertia of the symmetric and the positive-semidefinite approximations (in the Frobenius norm) of a low rank matrix.

Namely, our main result (Theorem 3.1.1) states that the positive and negative inertia indices of these approximations are upper bounded by the rank of the original matrix and this allows us to prove that the rank of the PSD approximation of a real matrix M is less than or equal to the rank of M. Using this result we will be able to combine Theorem 1.3.24 and 1.3.26 in order to propose a phylogenetic reconstruction method.

In this section, $M \in \mathcal{M}_n(\mathbb{R})$ is a real matrix and $S = \dfrac{M + M^T}{2}$ stands for its nearest symmetric matrix (see Theorem 1.2.11). The following result states that the positive and negative indices of inertia of S are upper bounded by $\text{rank}(M)$.

Theorem 3.1.1. *Let $M \in \mathcal{M}_n(\mathbb{R})$ be a matrix of $\text{rank}(M) = k$ and S the symmetric approximation of M. Then $i_+(S) \leq k$, $i_-(S) \leq k$ and $i_0(S) \geq \max\{0, n - 2k\}$.*

The following easy lemma is crucial to prove Theorem 3.1.1.

Lemma 3.1.2. *For any $v \in \mathbb{R}^n$ we have $v^T M v = v^T S v$.*

Proof. For any $v \in \mathbb{R}^n$, $v^T M v$ is a scalar, so $v^T M v = (v^T M v)^T$. It follows that

$$v^T S v = v^T \left(\frac{M + M^T}{2} \right) v = \frac{1}{2} \left(v^T M v + v^T M^T v \right) = \frac{1}{2} \left(v^T M v + (v^T M v)^T \right) =$$
$$= \frac{1}{2} \left(v^T M v + v^T M v \right) = v^T M v,$$

which proves the result. □

We are now ready to prove Theorem 3.1.1.

Proof of Theorem 3.1.1. We proceed by contradiction. Let (i_+, i_-, i_0) denote the inertia of S and suppose $i_+ = N$ with $N > k$.

Let $v_1, \dots, v_n$ be an orthonormal basis of eigenvectors of S (which exists by the Spectral Theorem) ordered so that $v_1, \dots, v_N$ correspond to the positive eigenvalues. That is, Sv_i equals $\lambda_i v_i$ for $i = 1, \cdots, n$, $\lambda_i > 0$ for all $i \in \{1, \dots, N\}$, and $\lambda_i \leq 0$ if $i > N$. Let V be the subspace spanned by these eigenvectors $V := \langle v_1, \dots, v_N \rangle$. Then for all non-zero $v \in V$ we have $v^T S v > 0$.

Grassman's formula states that $\dim(E + F) = \dim(E) + \dim(F) - \dim(E \cap F)$ if E, F are vector subspaces of the same vector space (see 4.4.19 in Meyer, 2000). Note that $\dim(V) = N$ and $\dim(\ker(M)) = n - k$ since $\text{rank}(M)$ equals k. Thus we get

$$\dim(V + \ker(M)) + \dim(V \cap \ker(M)) = \dim(V) + \dim(\ker(M)) = N + n - k. \quad (3.1)$$

Since $\dim(V + \ker(M)) \leq n$ and $N > k$, expression (3.1) gives us that $\dim(V \cap \ker(M)) \geq 1$. Consequently, there exists at least one element $w \neq \vec{0}$ such that $w \in V \cap \ker(M)$. By Lemma 3.1.2, $w^T M w$ equals $w^T S w$, which is positive since w belongs to V. But at the same time w belongs to $\ker(M)$ and hence $w^T M w = w^T \vec{0}$, which is equal to zero. This leads to a contradiction and we conclude $i_+ \leq k$.

An analogous argument can be used to prove that $i_- \leq k$. The fact that $i_0 \geq \max\{0, n-2k\}$ follows trivially since $i_+ + i_- + i_0 = n$. □

Using the previous result, it is easy to bound the rank of the PSD approximation of a matrix $M \in \mathcal{M}_n(\mathbb{R})$:

Corollary 3.1.3. *For any real matrix M, the rank of $psd(M)$ of a real matrix M is less than or equal to* $\operatorname{rank}(M)$.

Proof. There exists an orthonormal basis $v_1, \dots, v_n$ of eigenvectors of S, the symmetric approximation of M, with respective eigenvalues $\{\lambda_1, \dots, \lambda_n\}$ such that $S = P\Lambda P^T$ where $\Lambda = \operatorname{diag}(\lambda_i)$ and P is the orthogonal matrix with these eigenvectors as columns. In particular, $P^{-1} = P^T$. We denote by Λ^+ the diagonal matrix with entries $\lambda_i^+ = \max\{0, \lambda_i\}$.

We claim that $P\Lambda^+P^T$ is equal to the PSD approximation of M. In order to show this, first consider $\sigma_i = \frac{\lambda_i}{|\lambda_i|}$ if $\lambda_i \neq 0$ or $\sigma_i = 1$ if $\lambda_i = 0$, and define $\Sigma = \operatorname{diag}(\sigma_i)$ and $|\Lambda| = \operatorname{diag}(|\lambda_i|)$. Then it is immediate that $\Lambda = \Sigma|\Lambda|$ so $S = P\Sigma|\Lambda|P^T = \left(P\Sigma P^T\right)\left(P|\Lambda|P^T\right)$ and the last equality holds because P is orthogonal. Note that $U = P\Sigma P^T$ is an orthogonal matrix and $H = P|\Lambda|P^T$ is symmetric and positive-semidefinite, so $S = UH$ is a polar decomposition of S.

Note also that $(\sigma_i + 1)|\lambda_i| = \lambda_i + |\lambda_i| = 2\lambda_i^+$ and then $\frac{(\Sigma + Id)|\Lambda|}{2}$ is equal to Λ^+. Therefore, by Theorem 1.2.12 (see Lemma 2.4 on Higham, 1988) the PSD approximation $psd(M)$ of M is given by

$$\begin{aligned} psd(M) &= \frac{S+H}{2} = \frac{(P\Sigma P^T)(P|\Lambda|P^T) + P|\Lambda|P^T}{2} = \frac{(P\Sigma P^T + Id)P|\Lambda|P^T}{2} \\ &= \frac{P(\Sigma + Id)|\Lambda|P^T}{2} = P\Lambda^+P^T. \end{aligned}$$

From the new expression of $psd(M) = P\Lambda^+P^T$ we see that $\operatorname{rank}(psd(M)) = \#\{\lambda_i^+ \mid \lambda_i^+ \neq 0\}$ (λ_i^+ counted with multiplicity) which coincides with the positive inertia index of S. Finally since $i_+(S) \leq \operatorname{rank}(M)$ by Theorem 3.1.1, we obtain $\operatorname{rank}(psd(M)) \leq \operatorname{rank}(M)$. □

3.2 LEAF-TRANSFORMATIONS

In this section we introduce the leaf-transformations that correspond to a Markov action on a tensor p according to a split $A|B$. If p has arisen from a quartet T with certain substitution parameters, then an $A|B$ leaf-transformation on p produces a different tensor q that can be seen as a tensor arising on the same tree T with different parameters at the exterior edges.

Notation 3.2.1. Let p be a 4-tensor and let $i, j \in [4]$ with $i < j$. We denote by $N_{i,j}$ the matrix consisting on the double marginalization of p on the coordinates different from i and j, and with rows and columns labeled by the states at i and j, respectively. We call $N_{j,i}$ the transpose of $N_{i,j}$. For example, if $i = 1$ and $j = 3$, $(N_{13})_{xy} = \sum_{a,b} p_{x,a,y,b}$ and $(N_{31})_{xy} = \sum_{a,b} p_{y,a,x,b}$ and equivalently $N_{1,3} = p_{.+.+}$ and $N_{3,1} = (p_{.+.+})^t$.

When p is a distribution arising on some tree T with certain substitution parameters, the following lemma describes each $N_{i,j}$ in terms of these substitution parameters.

Lemma 3.2.2. *Let $p = \phi_T\left(\pi; \{M_e\}_{e\in E(T)}\right)$ be a distribution arising from a 4-leaf tree $T = T_{A|B}$ and let $i, j \in [4]$ with $i < j$. Assume the root of T is placed at the parent node of the leaves labeled by A. Then*

$$N_{i,j} = \begin{cases} M_i^t diag(\pi) M_5 M_j & \text{if } i \in A \text{ and } j \in B \\ M_i^t \left(diag(\pi) M_5\right)^t M_j & \text{if } i \in B \text{ and } j \in A \\ M_i^t diag(\pi) M_j & \text{if } \{i, j\} = A \\ M_i^t diag(M_5^t \pi) M_j & \text{if } \{i, j\} = B \end{cases} \tag{3.2}$$

Proof. The first and second cases follow from Lemma 1.3.16, where the statement is written for $T = T_{12|34}$ but can be adapted to any $T \in \mathcal{T}_4$.

To prove the other two cases we fix $T = T_{12|34}$, the proof for the other quartets is analogous. In order to compute $N_{1,2} = p_{..++} = (p_{...+})_{..+}$, we marginalize first over the fourth leaf and then over the third one. By Equation (1.27), $p_{...+} = \phi_{T_3}(\pi, \{M_1, M_2, M_5 M_3\})$ where T_3 is the trivalent 3-leaf tree. Lemma 1.3.15 implies that $p_{..++} = M_1^t \text{diag}(\pi) M_2$.

To prove the last case, we relocate the root of the tree according to Example 1.3.5. If $\pi' = M_5^t \pi$ and $M_5' = diag(\pi')^{-1} M_5^t diag(\pi)$ and the root is located at the parent node of leaves 3 and 4 it is straightforward to see that $p_{+...} = \phi_{T_3}\left(\pi'; M_5' M_2, M_3, M_4\right)$. Therefore, from Lemma 1.3.15 we have that $p_{++..} = (p_{+...})_{+..} = M_3^t diag(\pi') M_4 = M_3^t diag(M_5^t \pi) M_4$, which proves the last case. □

Lemma 3.2.3. *Let p be a 4-tensor and let $i, j, k \in [4]$ with $i < j, k$. Then, if $N_{i,j}$ is invertible*

$$N_{i,j}^{-1} N_{i,k} \in \mathbb{M}_4^*. \tag{3.3}$$

Proof. Let D_i denote the diagonal matrix whose entries are the triple marginalization of p on the coordinates different from i (for instance $D_1 = p_{.+++}$). Then the diagonal entries of D_i are the sums of rows of $N_{i,j}$ and $N_{i,k}$, that is $N_{i,j}\mathbf{1} = N_{i,k}\mathbf{1} = D_i$. Therefore, $N_{i,j}^{-1} N_{i,k}\mathbf{1} = N_{i,j}^{-1} D_i = \mathbf{1}$. □

As a consequence of the previous lemma, we can define the following Markov action.

Definition 3.2.4. Let p be a 4-tensor, then for any non-trivial split $A|B$ we define the *$A|B$ leaf-transformations* on p as the set of Markov actions:

$$p \longmapsto \left(p *_a N_{\tilde{b},a}^{-1} N_{\tilde{b},a^c}\right) *_b N_{\tilde{a},b}^{-1} N_{\tilde{a},b^c}\,, \tag{3.4}$$

where $a, \tilde{a} \in A$, $b, \tilde{b} \in B$, a^c denotes the complementary of a in A (respectively for b^c) and $N_{a,b}$ are the matrices defined above. Note that a, $\tilde{a}$ may coincide and the same for b, $\tilde{b}$. Note also the different role played by $\{a, b\}$ and by $\{\tilde{a}, \tilde{b}\}$ in the definition of the new tensor. Consider these 16 leaf-transformations following a lexicographic order on the words $ab\tilde{a}\tilde{b}$. Then, we denote by $p_k^{A|B}$ the resulting tensor obtained by applying the k-th $A|B$ leaf-transformation on p where $k \in [16]$.

Example 3.2.5. For instance, take the split $A|B$, with $A = \{1, 2\}$ and $B = \{3, 4\}$, then $p_1^{12|34}$ is the resulting tensor obtained by applying the Markov action with $a = \tilde{a} = 1$, $b = \tilde{b} = 3$, $a^c = 2$, and $b^c = 4$,

$$p \longmapsto p_1^{12|34} := \left(p *_1 N_{3,1}^{-1} N_{3,2}\right) *_3 N_{1,3}^{-1} N_{1,4} = \left(N_{3,1}^{-1} N_{3,2} \otimes Id \otimes N_{1,3}^{-1} N_{1,4} \otimes Id\right) \cdot p.$$

If $p = \phi_T(\pi; M_1, M_2, M_3, M_4, M_5)$ with $T = T_{12|34}$ then

$$p_1^{12|34} = \left(M_1^{-1} M_2 \otimes Id \otimes M_3^{-1} M_4 \otimes Id\right) \cdot p.$$

The following proposition is written for $T = T_{12|34}$ for simplicity, however there is an analogous statement for $T = T_{13|24}$ and $T = T_{14|23}$.

Proposition 3.2.6. *Let $p = \phi_T(\pi; M_1, M_2, M_3, M_4, M_5)$ be a distribution arising from the tree $T = T_{12|34}$. Let $p_k^{12|34}$ be the tensor obtained by applying a $12|34$ leaf-transformation*

on p corresponding to some $a \in \{1,2\}$ and $b \in \{3,4\}$. Then, we have

$$p_k^{12|34} = \phi_{12|34}(\pi, M_{a^c}, M_{a^c}, M_{b^c}, M_{b^c}, M_5),$$

where a^c (respectively b^c) is the complementary of a on $\{1,2\}$ (respectively b on $\{3,4\}$).

Proof. From Lemma 3.2.2 we have that $N_{a,b} = M_a^t diag(\pi) M_5 M_b$, if $a < b$. Therefore, assuming $a < \tilde{b}$,

$$N_{\tilde{b},a}^{-1} N_{\tilde{b},a^c} = \left(M_{\tilde{b}}^t M_5^t diag(\pi) M_a\right)^{-1} \left(M_{\tilde{b}}^t M_5^t diag(\pi) M_{a^c}\right) = M_a^{-1} M_{a^c} \tag{3.5}$$

and equivalently $N_{\tilde{a},b}^{-1} N_{\tilde{a},b^c} = M_b^{-1} M_{b^c}$. By Lemma 1.3.14 the tensor $p *_a M_a^{-1} M_{a^c}$ is the joint distribution arising from the same tree as p except that M_a has been replaced by M_{a^c}. And hence $\left(p *_a M_a^{-1} M_{a^c}\right) *_b M_b^{-1} M_{b^c}$ is the joint distribution arising from the same tree as p except that M_a has been replaced by M_{a^c}, and M_b by M_{b^c}. □

Therefore, the effect of the $A|B$ leaf-transformations on a distribution p that arises from the tree $T = T_{A|B}$ is that of replacing some of the external matrices so that both leaves in the same side of the tree share the same matrix, see Figure 3.1 as an illustration. Moreover, we infer two important conclusions from these last results.

Firstly, the 16 $A|B$ leaf-transformations on $p \in \operatorname{Im} \phi_T$ produce only 4 different points on $\mathcal{V}_T$ if $T = T_{A|B}$. Indeed, there are only 4 possible combinations if sister leaves have to share the same matrix. (see Figure 3.1 for the case $T = T_{12|34}$). Therefore, the 16 transformations correspond to four possible ways of computing each of these transformed distributions. However, these four computations do not produce the same tensor when applied to a distribution $\tilde{p}$ that has not arisen as a Markov process on T. In general, we have 16 $A|B$ leaf-transformations that can be applied to any distribution $\tilde{p} \in \mathbb{R}^{4^4}$.

The second conclusion is that if p lies on the variety $\mathcal{V}_T$ then $p_k^{A|B}$ also lies on the same variety. This can be generalized to any $A'|B'$ leaf-transformation as follows:

Proposition 3.2.7. *Let $p = \phi_T\left(\pi; \{M_e\}_{e \in E(T)}\right)$ be a distribution arising on a quartet T and $p_k^{A|B}$ a tensor obtained by applying an $A|B$ leaf-transformation on p (for any split $A|B$). Then*

$$p_k^{A|B} \in Im\, \phi_T.$$

Proof. Because of Lemma 1.3.13 $p_k^{A|B} = \phi_T(\pi; \tilde{M}_1, \tilde{M}_2, \tilde{M}_3, \tilde{M}_4, M_5)$ for some matrices $\tilde{M}_i$. Moreover, each $\tilde{M}_i \in \mathbb{M}_4^*$ by Lemma 3.2.3. □

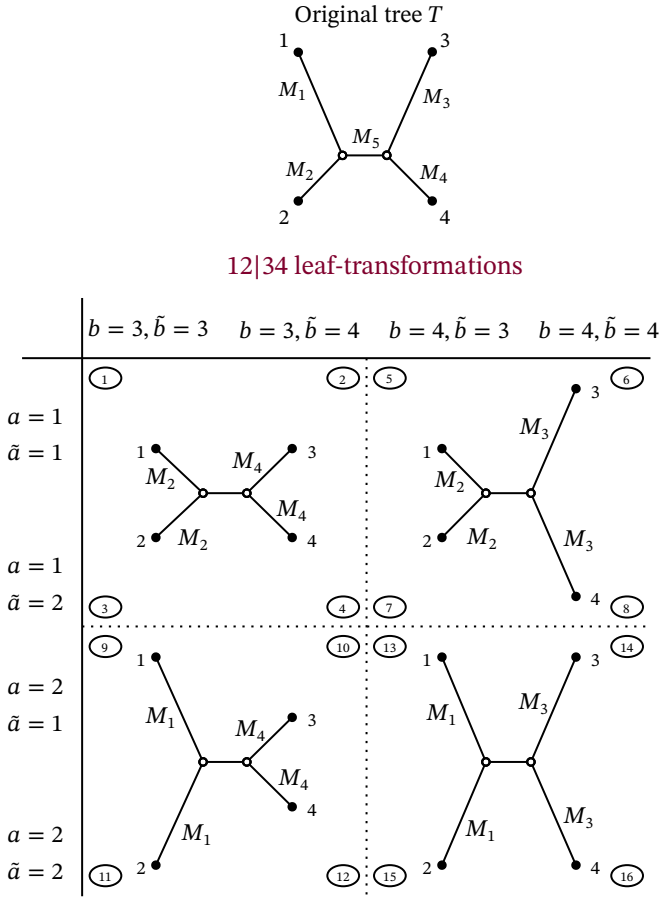


FIGURE 3.1: Markov matrices associated with the 12|34 leaf-transformations on a tensor p arising from $T_{12|34}$. Each tree corresponds to the four tensors $p_k^{12|34} = \left(p *_a N^{-1}_{\tilde{b},a} N_{\tilde{b},a^c}\right) *_b N^{-1}_{\tilde{a},b} N_{\tilde{a},b^c}$ producing the same tree. The rounded numbers denote the corresponding k for the values of $a, \tilde{a}$ in its row and $b, \tilde{b}$ in its column.

The proposition above tells us that if p is a distribution that arises from a Markov process on T, then for any split $A|B$ (being induced or not by T), the $A|B$ leaf-transformations applied to p produce 16 tensors $p_k^{A|B}$ that have also arisen on T but with different matrices at the exterior edges (see Figure 3.1 for the leaf-transformations corresponding to a split induced by T and Figure 3.2 and Figure 3.3 for leaf-transformations according to incompatible splits). Moreover, while the $A|B$ leaf-transformations applied to a distribution p produce *distributions* if $A|B$ is induced by T, the $A'|B'$ leaf-transformations (for splits not induced by T) may produce non-stochastic tensors since the new parameters may not be stochastic matrices. In Figure 3.2, for example, the transition matrices $M_5^{-1}M_1$ or $M_5^{-1}M_2$ on the 13-th 13|24 leaf-transformation might not be stochastic. In all cases, the transition matrix at the interior edge remains equal to the matrix M_5 of the original tree.

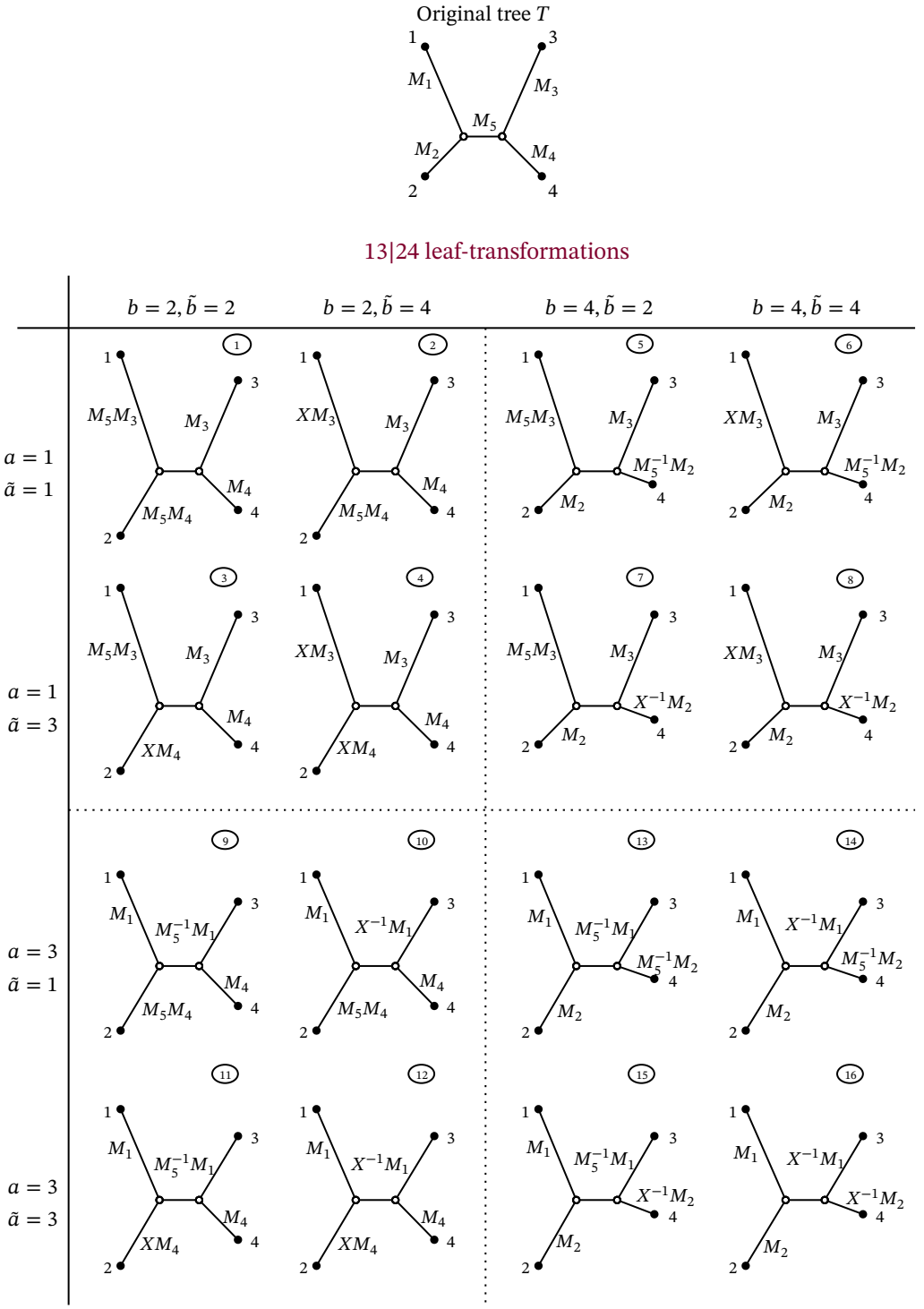
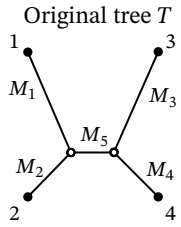


FIGURE 3.2: External matrices associated to the 13|24 leaf-transformations on a tensor p arising from the tree $T_{12|34}$. Each tree corresponds to the tensor $p_k^{13|24} = \left(p *_a N_{\tilde{b},a}^{-1} N_{\tilde{b},a^c}\right) *_b N_{\tilde{a},b}^{-1} N_{\tilde{a},b^c}$. The rounded number near each tree corresponds to k. The matrix at the interior edge is always M_5 and $X = \left(M_5^t diag(\pi)\right)^{-1} diag(M_5^t \pi)$.



14|23 leaf-transformations

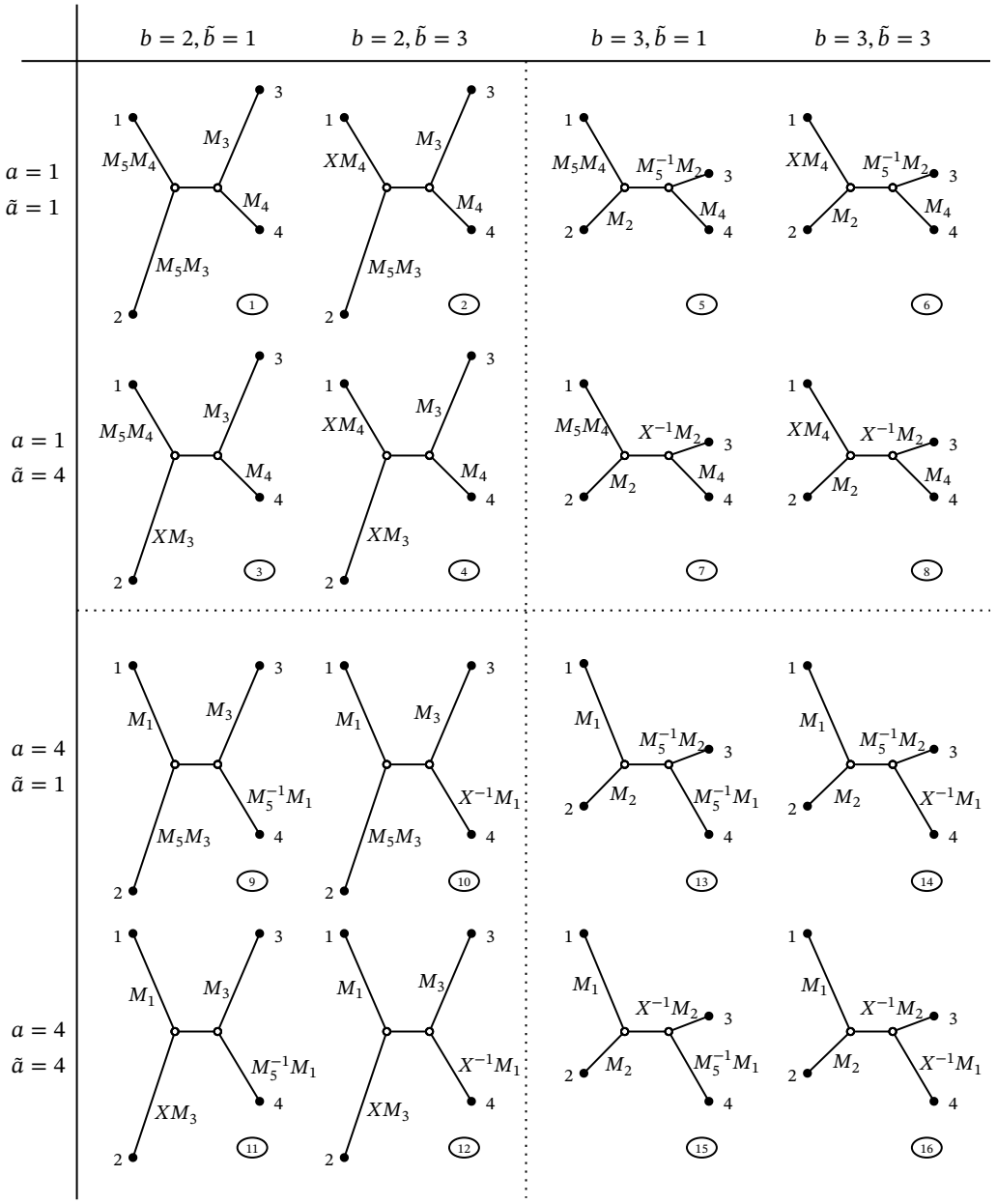


FIGURE 3.3: External matrices associated to the 13|24 leaf-transformations on a tensor p arising from the tree $T_{12|34}$ and following the notation of Figure 3.2 with $A = \{1, 4\}$ and $B = \{2, 3\}$.

Example 3.2.8. Let $p = \phi_T(\pi; M_1, M_2, M_3, M_4, M_5)$ be a distribution arising on $T = T_{12|34}$ and consider $A|B$ leaf-transformations $\left(p *_a N^{-1}_{\tilde{b},a} N_{\tilde{b},a^c}\right) *_b N^{-1}_{\tilde{a},b} N_{\tilde{a},b^c}$ with $A = \{1,3\}$ and $B = \{2,4\}$ as follows. The action over leaf 1 is done through the matrices

$$N^{-1}_{2,1} N_{2,3} = \left(M_2^t diag(\pi) M_1\right)^{-1} M_2^t diag(\pi) M_5 M_3 = M_1^{-1} M_5 M_3,$$
$$N^{-1}_{4,1} N_{4,3} = \left(M_4^t M_5^t diag(\pi) M_1\right)^{-1} M_4^t diag(M_5^t \pi) M_3 = M_1^{-1} X M_3,$$

where $X := \left(M_5^t diag(\pi)\right)^{-1} diag(M_5^t \pi)$, and through their inverses on the leaf 3. Similarly, the action over the leaves $2, 4$ is done through the matrices

$$N^{-1}_{1,2} N_{1,4} = \left(M_1^t diag(\pi) M_2\right)^{-1} M_1^t diag(\pi) M_5 M_4 = M_2^{-1} M_5 M_4;$$
$$N^{-1}_{3,2} N_{3,4} = \left(M_3^t M_5^t diag(\pi) M_2\right)^{-1} M_3^t diag(M_5^t \pi) M_4 = M_2^{-1} X M_4;$$

for $b = 2$, and their inverses for $b = 4$. The equalities above follow from Lemma 3.2.2 with $A = \{1,2\}$ and $B = \{3,4\}$. In Figure 3.2 we illustrate the effect of these sixteen 13|24 leaf-transformations on p.

In the remaining of this section we study the rank and the positive semidefiniteness of the flattening matrices of these leaf-transformations.

Lemma 3.2.9. *Let $p_k^{12|34}$ be a tensor obtained by applying a 12|34 leaf-transformation to a tensor p arising on the tree $T = T_{12|34}$, i.e. $p = \phi_{T_{12|34}}(\pi; M_1, M_2, M_3, M_4, M_5)$, where M_i are stochastic transition matrices. Then, the following properties are satisfied:*

(a) The flattening matrix $Flat_{12|34}\left(p_k^{12|34}\right)$ has rank at most 4.

(b) The flattening matrices $Flat_{13|24}\left(p_k^{12|34}\right)$ and $Flat_{14|23}\left(p_k^{12|34}\right)$ are positive semi-definite and have rank 16 for positive and non-singular parameters. Moreover, they are equal and can be written as the product

$$\left(M_{a^c} \otimes M_{a^c}\right)^t D \left(M_{b^c} \otimes M_{b^c}\right),$$

for the corresponding a^c and b^c and where D is the diagonal matrix containing the entries of $diag(\pi)M_5$.

Proof. These statements are consequences of Theorem 1.3.24, Remark 1.3.25, Theorem 1.3.26 and Corollary 1.3.27. □

Corollary 3.2.10. *Keeping the notation and hypotheses of the previous lemma we have:*

(a) The closest positive-semidefinite matrix to $Flat_{12|34}\left(p_k^{12|34}\right)$ *has rank at most* 4*; in other words,*

$$\delta_4\left(psd\left(Flat_{12|34}\left(p_k^{12|34}\right)\right)\right) = 0.$$

(b) $\delta_4\left(psd\left(Flat_{13|24}\left(p_k^{12|34}\right)\right)\right) = \delta_4\left(psd\left(Flat_{14|23}\left(p_k^{12|34}\right)\right)\right)$ *and is strictly positive for positive non-singular parameters.*

Proof. (a) From Corollary 3.1.3 we have that

$$\operatorname{rank}\left(psd\left(Flat_{12|34}\left(p_k^{12|34}\right)\right)\right) \leq \operatorname{rank}\left(Flat_{12|34}\left(p_k^{12|34}\right)\right)$$

and this rank is at most 4 by Lemma 3.2.9 (a). Therefore, it is clear that $\delta_4\left(psd\left(Flat_{12|34}\left(p_k^{12|34}\right)\right)\right) = 0$.

(b) By Lemma 3.2.9 (b) we have that

$$psd\left(Flat_{13|24}\left(p_k^{12|34}\right)\right) = Flat_{13|24}\left(p_k^{12|34}\right) = Flat_{14|23}\left(p_k^{12|34}\right) = psd\left(Flat_{14|23}\left(p_k^{12|34}\right)\right).$$

Moreover they have rank 16 for positive non-singular parameters (see Remark 1.3.25) and the claim follows. □

Lemma 3.2.11. *Let* $p_k^{13|24}$ *be the tensor obtained by applying a* $13|24$ *leaf-transformation to a tensor* $p = \phi_T(\pi; M_1, \ldots, M_5)$ *arising from the tree* $T = T_{12|34}$*. Then the following properties are satisfied:*

(a) The flattening matrix $Flat_{12|34}\left(p_k^{13|24}\right)$ *has rank at most* 4*.*

(b) If M_i *are positive K81 matrices for* $i = 1, \cdots, 5$ *and* π *is uniform, then*

$$\delta_4\left(psd\left(Flat_{13|24}\left(p_k^{13|24}\right)\right)\right) > 0.$$

Proof. (a) By Proposition 3.2.7 the tensor $p_k^{13|24}$ lies in $\mathcal{V}_{T_{12|34}}$ and therefore the flattening $Flat_{12|34}\left(p_k^{13|24}\right)$ has rank less than or equal to 4 by Theorem 1.3.24.

(b) Recall that any $13|24$ leaf-transformation on p produces a tensor $p_k^{13|24}$ that has arisen on T with modified transition matrices at the exterior edges as illustrated in Figure 3.2. We follow the notation of that figure in this proof. The flattening matrix $Flat_{13|24}\left(p_k^{13|24}\right)$ can be written in terms of these new matrices (see Theorem 1.3.22 and Lemma 1.3.23) as:

$$Flat_{13|24}\left(p_k^{13|24}\right) = (AM_i \otimes BM_i)^t \; D \; (CM_j \otimes DM_j)$$

where D is the 16×16 diagonal matrix with the entries of $\frac{1}{4}M_5$, and M_i, M_j, A, B, C, D depend on the given leaf-transformation (see notation in Figure 3.2). Since M_5 is a K81 matrix and π is uniform, we get

$$\begin{aligned} X &= \left(M_5^t diag(\pi)\right)^{-1} diag(M_5^t \pi) = \left(M_5 diag(\pi)\right)^{-1} diag(M_5 \pi) \\ &= \left(\frac{1}{4}M_5\right)^{-1} \frac{1}{4} Id = M_5^{-1}, \end{aligned}$$

and therefore A, B, C, D are either equal to M_5, to M_5^{-1} or to the identity matrix. For instance, if $k = 1$, $M_i = M_3$, $M_j = M_4$, $A = C = M_5$ and $B = D$ and are equal to the identity matrix, see Figure 3.2.

As M_i, A, B, C and D are K81 matrices they diagonalize through the Hadamard matrix H and for each matrix the multiplicity of 1 as an eigenvalue is exactly one if the matrices have positive entries (see Lemma 1.1.28 and Theorem 1.2.15). Write $\{\lambda_k^i\}_{k=\mathtt{C},\mathtt{G},\mathtt{T}}$ for the three eigenvalues of the matrix M_i other than 1. And denote by $\{\alpha_k\}_{k=\mathtt{C},\mathtt{G},\mathtt{T}}$, $\{\beta_k\}_{k=\mathtt{C},\mathtt{G},\mathtt{T}}$, $\{\gamma_k\}_{k=\mathtt{C},\mathtt{G},\mathtt{T}}$ and $\{\delta_k\}_{k=\mathtt{C},\mathtt{G},\mathtt{T}}$ the eigenvalues of A, B, C and D (respectively) different from one. We can assume that all these eigenvalues are positive since the transition matrices should not be too far from the identity matrix.

As we have seen in the proof of Corollary 3.1.3, for any matrix M the rank of the nearest positive semidefinite matrix $psd(M)$ is equal to the number of positive eigenvalues of the symmetric matrix $\frac{1}{2}\left(M + M^t\right)$. Thus, it is enough to show that the symmetric matrix

$$S_k := \frac{1}{2}\left(Flat_{13|24}\left(p_k^{13|24}\right) + Flat_{13|24}\left(p_k^{13|24}\right)^t\right)$$

has at least 5 positive eigenvalues.

Write $\overline{S}_k := (H \otimes H)^t S_k (H \otimes H)$. Because of Sylvester's law of inertia (see Theorem 1.2.5), both matrices, S_k and $\overline{S}_k$ have the same number of positive and negative eigenvalues. To prove that $\overline{S}_k$ has at least 5 positive eigenvalues we shall use Theorem 1.2.7 and see that the first five leading principal minors of $\overline{S}_k$ are positive.

Consider the 5×5 submatrix of $\overline{S}_{k,5}$ obtained by removing the last 11 rows and columns of $\overline{S}_k$. It can be written in terms of the eigenvalues of the matrices M_i, M_j, M_5, A, B, C and D as:

$$\overline{S}_{k,5} = \frac{1}{4^4}\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \beta_C\delta_C\lambda_C^i\lambda_C^j & 0 & 0 & \frac{1}{2}(\beta_C\gamma_C+\alpha_C\delta_C)\lambda_C^5\lambda_C^i\lambda_C^j \\ 0 & 0 & \beta_G\delta_G\lambda_G^i\lambda_G^j & 0 & 0 \\ 0 & 0 & 0 & \beta_T\delta_T\lambda_T^i\lambda_T^j & 0 \\ 0 & \frac{1}{2}(\beta_C\gamma_C+\alpha_C\delta_C)\lambda_C^5\lambda_C^i\lambda_C^j & 0 & 0 & \alpha_C\gamma_C\lambda_C^i\lambda_C^j \end{pmatrix}.$$

The leading principal minors of $\overline{S}_k$ are equal to the first five leading principal minors of $\overline{S}_{k,5}$:

$$\Delta_1 = 1, \quad \Delta_2 = \beta_C\delta_C\lambda_C^i\lambda_C^j, \quad \Delta_3 = \beta_C\delta_C\lambda_C^i\lambda_C^j\beta_G\delta_G\lambda_G^i\lambda_G^j = \Delta_2\delta_G\lambda_G^i\lambda_G^j,$$

$$\Delta_4 = \beta_C\delta_C\lambda_C^i\lambda_C^j\beta_G\delta_G\lambda_G^i\lambda_G^j\beta_T\delta_T\lambda_T^i\lambda_T^j = \Delta_2\Delta_3\beta_T\delta_T\lambda_T^i\lambda_T^j,$$

$$\Delta_5 = -\frac{1}{4}\left(\left(\lambda_C^5\right)^2(\beta_C\gamma_C+\alpha_C\delta_C)^2 - 4\alpha_C\beta_C\gamma_C\delta_C\right)\cdot\beta_G^2\delta_G^2\left(\lambda_C^i\right)^2\lambda_G^i\lambda_T^i\left(\lambda_C^j\right)^2\lambda_G^j\lambda_T^j.$$

We have that $\Delta_i > 0$ $(i = 1, 2, 3, 4)$ because all the eigenvalues are positive. In order to see that Δ_5 is also positive we only need to prove that the factor $\Delta_5' = -\left(\lambda_C^5\right)^2(\beta_C\gamma_C+\alpha_C\delta_C)^2 + 4\alpha_C\beta_C\gamma_C\delta_C$ is positive.

We prove it for all possible values of A, B, C and D. According to Definition 3.2.4 (see Figure 3.2) we have these different possibilities:

- Suppose that the set $\{A, B, C, D\}$ equals the set $\{M_5, M_5, Id, Id\}$, i.e., that there are two matrices equal to M_5 and two equal to the identity. Then, since $(\lambda_C^5)^2 < 1$,

$$\Delta_5' = \begin{cases} -\left(\lambda_C^5\right)^2\left((\lambda_C^5)^2+1\right)^2 + 4\left(\lambda_C^5\right)^2 = \left(\lambda_C^5\right)^2\left(-\left((\lambda_C^5)^2+1\right)^2+4\right) > 0 & \text{or} \\ -\left(\lambda_C^5\right)^2\left(\lambda_C^5+\lambda_C^5\right)^2 + 4\left(\lambda_C^5\right)^2 = \left(\lambda_C^5\right)^2\left(-4\left(\lambda_C^5\right)^2+4\right) > 0. \end{cases}$$

- Suppose $\{A, B, C, D\} = \{M_5, M_5^{-1}, Id, Id\}$, then

$$\Delta_5' = \begin{cases} -\left(\lambda_C^5\right)^2(1+1)^2 + 4 = 4 - 4\left(\lambda_C^5\right)^2 > 0 & \text{or} \\ -\left(\lambda_C^5\right)^2\left(\lambda_C^5+\dfrac{1}{\lambda_C^5}\right)^2 + 4 = 4 - \left(\left(\lambda_C^5\right)^2+1\right)^2 > 0. \end{cases}$$

- Finally, suppose $\{A,B,C,D\} = \{M_5^{-1}, M_5^{-1}, Id, Id\}$, then

$$\Delta_5' = \begin{cases} -\left(\lambda_C^5\right)^2\left(\left(\frac{1}{\lambda_C^5}\right)^2+1\right)^2 + 4\left(\frac{1}{\lambda_C^5}\right)^2 = \left(\frac{1}{\lambda_C^5}\right)^2\left(4-\left(1+\left(\lambda_C^5\right)^2\right)^2\right) > 0 \quad \text{or} \\ -\left(\lambda_C^5\right)^2\left(\frac{1}{\lambda_C^5}+\frac{1}{\lambda_C^5}\right)^2 + 4\left(\frac{1}{\lambda_C^5}\right)^2 = \left(\frac{1}{\lambda_C^5}\right)^2\left(4-2\left(\lambda_C^5\right)^2\right) > 0. \end{cases}$$

We conclude that the matrix $\overline{S}_k$ has at least 5 positive eigenvalues and so does S_k. This implies that $rk\left(psd\left(Flat_{13|24}\left(p_1^{13|24}\right)\right)\right)$ is greater than or equal to 5 and therefore, $\delta_4\left(psd\left(Flat_{13|24}\left(p_1^{13|24}\right)\right)\right) > 0$.

□

The last issue in this section is the symmetry of the flattening matrices of an $A|B$ leaf-transformation.

Remark 3.2.12. Given any tensor $p \in \mathbb{R}^{4^4}$ we can compute, a priori, 8 different flattenings of p regarding any split $A|B$. Namely, suppose $A = \{1,2\}$ and $B = \{3,4\}$ then we have

$$\begin{array}{cccc} Flat_{12|34}(p), & Flat_{12|43}(p), & Flat_{21|34}(p), & Flat_{21|43}(p), \\ Flat_{34|12}(p), & Flat_{43|12}(p), & Flat_{34|21}(p), & Flat_{43|21}(p). \end{array} \tag{3.6}$$

However there are many symmetries between these matrices. For instance $Flat_{aa'|bb'}(p) = Flat_{bb'|aa'}(p)^t$, for any a, a', b, b', which implies that these two matrices have the same eigenvalues, singular values, and therefore the same rank. Moreover they have the same closest symmetric and PSD matrix. Another relation is the following

$$Flat_{aa'|bb'}(p) = P^t Flat_{a'a|b'b}(p) P$$

where P is a permutation matrix. Then these two matrices share the same singular values. Moreover $psd(Flat_{aa'|bb'}(p)) = P^t\, psd(Flat_{a'a|b'b}(p))P$ and therefore $psd(Flat_{aa'|bb'}(p))$ and $psd(Flat_{a'a|b'b}(p))$ have the same singular values.

Leaf-transformations on mixtures

Leaf-transformations can also be applied to distributions p that have arisen from mixtures of distributions on trees, although their effect on p is harder to interpret. In this section we present the effect of leaf-transformations on two types of mixtures and we study properties of the flattenings of these transformations.

Remark 3.2.13. Consider a mixture distribution $p = \lambda_1 p_1 + \lambda_2 p_2$ where p_1 and p_2 are distributions that have arisen on T with certain substitution parameters and $\lambda_1 + \lambda_2 = 1$. Observe that for any $i \in [4]$ and any $N \in \mathbb{M}_4^*$, if $\mathbf{1} = (1,1,1,1)^t$,

$$p *_i \mathbf{1} = \lambda_1 p_1 *_i \mathbf{1} + \lambda_2 p_2 *_i \mathbf{1}, \quad \text{and} \quad p *_i N = \lambda_1 p_1 *_i N + \lambda_2 p_2 *_i N$$

because $*_i$ is a linear operation.

We need the following observation

Remark 3.2.14. Consider the quartet $T = T_{12|34}$, and the tensors $q_\pi = \phi_T(\pi; Id, Id, Id, Id, M)$ and $q = \phi_T(\mathbf{1}; Id, Id, Id, Id, M)$, $M \in \mathcal{M}_4(\mathbb{R})$ and $\pi \in \mathbb{R}^4$. Then

$$Flat_{12|34}(q_\pi) = \big(diag(\pi) \otimes Id\big)\, Flat_{12|34}(q),$$
$$Flat_{13|24}(q_\pi) = \big(diag(\pi) \otimes Id\big)\, Flat_{13|24}(q).$$

as can be deduced from Theorem 1.3.22.

The next result presents the first kind of mixture under study.

Proposition 3.2.15. *Consider $T = T_{12|34}$ and let $p_1 = \phi_T(\pi_1; M_1, M_2, M_3, M_4, M_5)$ and $p_2 = \phi_T(\pi_2; \tilde{M}_1, \tilde{M}_2, \tilde{M}_3, \tilde{M}_4, \tilde{M}_5)$ be two distributions arising on T with non-singular substitution parameters. Assume there exists a unique $i \in [5]$ such that $M_i \neq \tilde{M}_i$ and $M_j = \tilde{M}_j$ for all $j \neq i$. Consider the mixture $p = \lambda_1 p_1 + \lambda_2 p_2$ with $\lambda_1 + \lambda_2 = 1$. Then, for any k,*

$$rank\left(Flat_{12|34}\left(p_k^{12|34}\right)\right) \leq 4.$$

Moreover,

$$Flat_{13|24}\left(p_k^{12|34}\right) = Flat_{14|23}\left(p_k^{12|34}\right) \text{ are PSD matrices.}$$

Proof. In this proof we only consider $12|34$ leaf-transformations on p, and we will use p^k for the transformed $p_k^{12|34}$. We divide the proof in two cases.

Case $M_5 \neq \tilde{M}_5$:

Assume the matrix that differs between p_1 and p_2 is the corresponding to the interior edge: $p_1 = \phi_T(\pi_1; M_1, M_2, M_3, M_4, X_1)$ and $p_2 = \phi_T(\pi_2; M_1, M_2, M_3, M_4, X_2)$ with $X_1 \neq X_2$.

Let $a, \tilde{a} \in \{1,2\}$ and $b, \tilde{b} \in \{3,4\}$. For $l = 1,2$ denote by $N_{a,b}^l$ the double marginalization of p_l on coordinates different from a and b and write $N_{a,b}$ for the corresponding

marginalization on p. By Lemma 3.2.2,

$$N_{a,b} = \lambda_1 N^1_{a,b} + \lambda_2 N^2_{a,b} = M_a^t \left(\lambda_1 diag(\pi_1) X_1 + \lambda_2 diag(\pi_2) X_2\right) M_b.$$

Therefore $N^{-1}_{\tilde{b},a} N_{\tilde{b},a^c} = M_a^{-1} M_{a^c}$ and $N^{-1}_{\tilde{a},b} N_{\tilde{a},b^c} = M_b^{-1} M_{b^c}$. Using Remark 3.2.13 and Proposition 3.2.6

$$\begin{aligned} p_k &= \lambda_1 \left(p_1 *_a M_a^{-1} M_{a^c}\right) *_b M_b^{-1} M_{b^c} + \lambda_2 \left(p_2 *_a M_a^{-1} M_{a^c}\right) *_b M_b^{-1} M_{b^c} \\ &= \lambda_1 p_{1,k} + \lambda_2 p_{2,k} \end{aligned}$$

where $p_{l,k} := \phi_{12|34}(\pi_l, M_{a^c}, M_{a^c}, M_{b^c}, M_{b^c}, X_l)$ for $l = 1, 2$. Then we write the flattening

$$\begin{aligned} Flat_{12|34}(p_k) =& \lambda_1 Flat_{12|34}\left(p_{1,k}\right) + \lambda_2 Flat_{12|34}\left(p_{2,k}\right) = \\ =& \lambda_1 \left(M_{a^c} \otimes M_{b^c}\right)^t \left(Flat_{12|34}(q_1)\right) \left(M_{a^c} \otimes M_{b^c}\right) + \\ & \lambda_2 \left(M_{a^c} \otimes M_{b^c}\right)^t \left(Flat_{12|34}(q_2)\right) \left(M_{a^c} \otimes M_{b^c}\right) \\ =& \left(M_{a^c} \otimes M_{b^c}\right)^t \left(\lambda_1 Flat_{12|34}(q_1) + \lambda_2 Flat_{12|34}(q_2)\right) \left(M_{a^c} \otimes M_{b^c}\right), \end{aligned}$$

where q_1 and q_2 are the distributions arising on T with the identity matrix at the exterior edges and the matrices X_1 and X_2 at the interior edge (see Theorem 1.3.22). The matrix in the middle of the above expression has rank 4 (see the matrix in (1.32)). Similarly,

$$\begin{aligned} &Flat_{13|24}(p_k) = \lambda_1 Flat_{13|24}\left(p_{1,k}\right) + \lambda_2 Flat_{13|24}\left(p_{2,k}\right) = \\ &= \left(M_{a^c} \otimes M_{b^c}\right)^t \left(\lambda_1 D_1 + \lambda_2 D_2\right) \left(M_{a^c} \otimes M_{b^c}\right) \end{aligned}$$

where D_1 and D_2 are diagonal matrices with entries of $diag(\pi_1)X_1$ and $diag(\pi_2)X_2$, respectively. If the parameters are non-negative then $\lambda_1 D_1 + \lambda_2 D_2$ is a diagonal matrix with non-negative entries and $Flat_{13|24}(p_k)$ is PSD. On the other hand it is straighforward to see that $Flat_{13|24}(p_k) = Flat_{14|23}(p_k)$.

Case $M_i \neq \tilde{M}_i$ with $i \neq 5$:

We focus on the case $M_1 \neq \tilde{M}_1$, as the other cases are analogous. Denote $p_1 = \phi_T(\pi_1; X_1, M_2, M_3, M_4, M_5)$ and $p_2 = \phi_T(\pi_2; X_2, M_2, M_3, M_4, M_5)$ with $X_1 \neq X_2$.

The double marginalizations, for $b \in \{3, 4\}$ can be written as

$$\begin{aligned} N_{1,b} &= \left(\lambda_1 X_1^t diag(\pi_1) + \lambda_2 X_2^t diag(\pi_2)\right) M_5 M_b, \\ N_{2,b} &= M_2^t \left(\lambda_1 diag(\pi_1) + \lambda_2 diag(\pi_2)\right) M_5 M_b. \end{aligned} \tag{3.7}$$

We start considering only the first eight transformations: $p_k = \left(p *_1 N_{\tilde{b},1}^{-1} N_{\tilde{b},2}\right) *_b N_{\tilde{a},b}^{-1} N_{\tilde{a},b^c}$, for $k \in [8]$. According to (3.7) we have that

$$N_{\tilde{b},1}^{-1} N_{\tilde{b},2} = Y M_2, \quad N_{\tilde{a},b}^{-1} N_{\tilde{a},b^c} = M_b^{-1} M_{b^c},$$

where $Y := (\lambda_1 diag(\pi_1) X_1 + \lambda_2 diag(\pi_2) X_2)^{-1} (\lambda_1 diag(\pi_1) + \lambda_2 diag(\pi_2))$.

Analogously to the previous case, we have that for $k \in [8]$,

$$p_k = \lambda_1 p_{1,k} + \lambda_2 p_{2,k}$$

where $p_{l,k} := \phi_{12|34}(\pi_l, X_l Y M_2, M_2, M_{b^c}, M_{b^c}, M_5)$. We first write the flattening with respect to the split 12|34:

$$\begin{aligned}
&Flat_{12|34}(p_k) = \lambda_1 Flat_{12|34}(p_{1,k}) + \lambda_2 Flat_{12|34}(p_{2,k}) = \\
=&\lambda_1 (X_1 Y M_2 \otimes M_2)^t Flat_{12|34}(q_{\pi_1})(M_{b^c} \otimes M_{b^c}) + \\
&\lambda_2 (X_2 Y M_2 \otimes M_2)^t Flat_{12|34}(q_{\pi_2})(M_{b^c} \otimes M_{b^c}) = \\
=&\Big(\lambda_1 (X_1 Y M_2 \otimes M_2)^t (diag(\pi_1) \otimes Id) + \lambda_2 (X_2 Y M_2 \otimes M_2)^t (diag(\pi_2) \otimes Id)\Big) \cdot \\
&\cdot Flat_{12|34}(q)(M_{b^c} \otimes M_{b^c})
\end{aligned}$$

where $q_{\pi_i} = \phi_T(\pi_i; Id, Id, Id, Id, M_5)$ and $q = \phi_T(\mathbf{1}; Id, Id, Id, Id, M_5)$ and the last equality follows from Remark 3.2.14. Because of Theorem 1.3.24 rank $Flat_{12|34}(p_k) \leq 4$. On the other hand,

$$\begin{aligned}
&\lambda_l Flat_{13|24}(p_{l,k}) = \lambda_l (X_l Y M_2 \otimes M_{b^c})^t Flat_{13|24}(q_{\pi_l})(M_2 \otimes M_{b^c}) \\
&= (M_2 \otimes M_{b^c})^t \lambda_l (X_l Y \otimes Id)^t Flat_{13|24}(q_{\pi_l})(M_2 \otimes M_{b^c})
\end{aligned}$$

and then

$$\begin{aligned}
&Flat_{13|24}(p_k) = \lambda_1 Flat_{13|24}(p_{1,k}) + \lambda_2 Flat_{13|24}(p_{2,k}) = \\
&= (M_2 \otimes M_{b^c})^t \left(\lambda_1 (X_1 Y \otimes Id)^t Flat_{13|24}(q_{\pi_1}) + \lambda_2 (X_2 Y \otimes Id)^t Flat_{13|24}(q_{\pi_2})\right)(M_2 \otimes M_{b^c}) = \\
&= (M_2 \otimes M_{b^c})^t \left(\lambda_1 Flat_{13|24}(q_{\pi_1}) + \lambda_2 Flat_{13|24}(q_{\pi_2})\right)(M_2 \otimes M_{b^c}),
\end{aligned}$$

where the last equality follows from forthcoming Lemma 3.2.16. Then, if $(p_l)_k$ denotes the k-th transformation of p_l and the parameters are stochastic then $Flat_{13|24}(p_k) = \lambda_1 Flat_{13|24}((p_1)_k) + \lambda_2 Flat_{13|24}((p_2)_k)$ is a PSD matrix since it is the sum of two PSD matrices.

We now turn to the case $k = 9, \dots, 16$. We can proceed as above and get

$$p_k = \left(p *_2 N^{-1}_{\tilde{b},2} N_{\tilde{b},1}\right) *_b N^{-1}_{\tilde{a},b} N_{\tilde{a},b^c} = \lambda_1 p_{l,k} + \lambda_2 p_{l,k}$$

where $p_{l,k} = \phi_{12|34}\left(\pi_l, X_l, Y^{-1}, M_{b^c}, M_{b^c}, M_5\right)$. Proceeding as before andby Remark 3.2.14 we have

$$\begin{aligned} Flat_{12|34}(p_k) =& \Big(\lambda_1 \left(X_1 \otimes Y^{-1}\right)^t \left(diag(\pi_1 \otimes Id)\right) + \lambda_2 \left(X_2 \otimes Y^{-1}\right)^t \left(diag(\pi_2 \otimes M)\right)\Big) \cdot \\ & Flat_{12|34}(q)\left(M_{b^c} \otimes M_{b^c}\right) \end{aligned}$$

which, by the same argument as above, has rank at most 4.

Consider the 13|24 flattening of $p_{1,k}$:

$$\begin{aligned} &\lambda_1 Flat_{13|24}\left(p_{1,k}\right) = \\ &= \lambda_1 \left(X_1 \otimes M_{b^c}\right)^t Flat_{13|24}(q_{\pi_1})\left(Y^{-1} \otimes M_{b^c}\right) = \\ &= \left(Id \otimes M_{b^c}\right)^t \lambda_1 \left(X_1 \otimes Id\right)^t Flat_{13|24}(q_{\pi_1})\left(Y^{-1} \otimes Id\right)\left(Id \otimes M_{b^c}\right) = \\ &= \left(Id \otimes M_{b^c}\right)^t \lambda_1 \left(X_1 \otimes Id\right)^t \left(diag(\pi_1) \otimes Id\right) Flat_{13|24}(q)\left(Y^{-1} \otimes Id\right)\left(Id \otimes M_{b^c}\right) = \\ &= \left(Id \otimes M_{b^c}\right)^t \left(\lambda_1 X_1^t diag(\pi_1) \otimes Id\right) Flat_{13|24}(q)\left(Y^{-1} \otimes Id\right)\left(Id \otimes M_{b^c}\right) \end{aligned}$$

and then

$$\begin{aligned} &Flat_{13|24}(p_k) = \lambda_1 Flat_{13|24}\left(p_{1,k}\right) + \lambda_2 Flat_{13|24}\left(p_{2,k}\right) = \\ &= \left(Id \otimes M_{b^c}\right)^t \left(\left(\lambda_1 X_1^t diag(\pi_1) + \lambda_2 X_2^t diag(\pi_2)\right) \otimes Id\right) Flat_{13|24}(q)\left(Y^{-1} \otimes Id\right)\left(Id \otimes M_{b^c}\right). \end{aligned}$$

Note that $Y^{-1} = \left(\lambda_1 diag(\pi_1) + \lambda_2 diag(\pi_2)\right)^{-1}\left(\lambda_1 diag(\pi_1) X_1 + \lambda_2 diag(\pi_2) X_2\right)$ and write $U := \lambda_1 diag(\pi_1) X_1 + \lambda_2 diag(\pi_2) X_2$. Then,

$$Y^{-1} \otimes Id = \left(\left(\lambda_1 diag(\pi_1) + \lambda_2 diag(\pi_2)\right)^{-1} \otimes Id\right)\left(U \otimes Id\right),$$

and the flattening

$$Flat_{13|24}(p_k) = \left(U \otimes M_{b^c}\right)^t Flat_{13|24}(q)\left(\left(\lambda_1 diag(\pi_1) + \lambda_2 diag(\pi_2)\right)^{-1} \otimes Id\right)\left(U \otimes M_{b^c}\right).$$

Since $\lambda_1 diag(\pi_1) + \lambda_2 diag(\pi_2)$ is a diagonal matrix with positive entries at the diagonal so is its inverse. In conclusion $Flat_{13|24}(q)\left(\lambda_1 diag(\pi_1) + \lambda_2 diag(\pi_2)\right)^{-1}$ is diagonal with non-negative entries and then $Flat_{13|24}(p_k)$ is PSD. □

To conclude the proof of Proposition 3.2.15, we need the following lemma.

Lemma 3.2.16. *Keeping the above notation, we have*

$$\begin{aligned}\lambda_1 (X_1 Y \otimes Id)^t \, Flat_{13|24}(q_{\pi_1}) + \lambda_2 (X_2 Y \otimes Id)^t \, Flat_{13|24}(q_{\pi_2}) = \\ \lambda_1 Flat_{13|24}(q_{\pi_1}) + \lambda_2 Flat_{13|24}(q_{\pi_2}).\end{aligned} \tag{3.8}$$

In order to prove this equality we need a technical result.

Lemma 3.2.17 (Sherman–Morrison–Woodbury formula, see Meyer (2000))**.** *Let $A, B \in \mathcal{M}_n(\mathbb{R})$ be two invertible matrices and suppose $A + B$ is also invertible, then*

$$(A+B)^{-1} = A^{-1} - A^{-1}(AB^{-1} + Id)^{-1}.$$

In particular, if Z, $Z + Id$ and $Z - Id$ are invertible matrices, we have

$$(Z+Id)^{-1} + (Z^{-1} + Id)^{-1} = Id. \tag{3.9}$$

Proof of Lemma 3.2.16. By Remark 3.2.14 we have that

$$\begin{aligned}&\lambda_1 (X_1 Y \otimes Id)^t \, Flat_{13|24}(q_{\pi_1}) + \lambda_2 (X_2 Y \otimes Id)^t \, Flat_{13|24}(q_{\pi_2}) = \\ =&\lambda_1 (X_1 Y \otimes Id)^t \, (diag(\pi_1) \otimes Id) \, Flat_{13|24}(q) + \\ &\lambda_2 (X_2 Y \otimes Id)^t \, (diag(\pi_2) \otimes Id) \, Flat_{13|24}(q) = \\ =&\left(\left(\lambda_1 Y^t X_1^t diag(\pi_1) + \lambda_2 Y^t X_2^t diag(\pi_2)\right) \otimes Id \right) Flat_{13|24}(q)\end{aligned}$$

and

$$\lambda_1 Flat_{13|24}(q_{\pi_1}) + \lambda_2 Flat_{13|24}(q_{\pi_2}) = \left((\lambda_1 diag(\pi_l) + \lambda_2 diag(\pi_2)) \otimes Id \right) Flat_{13|24}(q).$$

Thus, it is enough to show that $\lambda_1 Y^t X_1^t diag(\pi_1) + \lambda_2 Y^t X_2^t diag(\pi_2) = \lambda_1 diag(\pi_l) + \lambda_2 diag(\pi_2)$ or equivalently that

$$\lambda_1 diag(\pi_1) X_1 Y + \lambda_2 diag(\pi_2) X_2 Y = \lambda_1 diag(\pi_l) + \lambda_2 diag(\pi_2). \tag{3.10}$$

Write $A_i := \lambda_i diag(\pi_i)$ and $Y = (A_1 X_1 + A_2 X_2)^{-1} (A_1 + A_2)$. We start computing the first term of (3.10) using Lemma 3.2.17.

$$\begin{aligned}&\lambda_1 diag(\pi_1) X_1 Y = A_1 X_1 \, (A_1 X_1 + A_2 X_2)^{-1} (A_1 + A_2) = \\ &= \left(Id - (Z + Id)^{-1}\right) (A_1 + A_2),\end{aligned}$$

where $Z = A_1X_1X_2^{-1}A_2^{-1}$. With an analogous computation we get $\lambda_2 diag(\pi_2)X_2Y = \left(Id-(Z^{-1}+Id)^{-1}\right)(A_1+A_2)$. Finally, a direct computation using (3.9) gives the equality (3.10). □

Proposition 3.2.18. *Consider $T = T_{12|34}$ and let $p_1 = \phi_T(\pi_1; M_1, M_1, N_1, N_1, X_1)$ and $p_2 = \phi_T(\pi_2; M_2, M_2, N_2, N_2, X_2)$ be two distributions arising on T with M_i, N_i and X_i stochastic matrices. Consider the mixture $p = \lambda_1 p_1 + \lambda_2 p_2$ with $\lambda_1 + \lambda_2 = 1$. Then,*

$$Flat_{13|24}\left(p_k^{12|34}\right) = Flat_{14|23}\left(p_k^{12|34}\right) \text{ are PSD matrices.}$$

Proof. In this case, for any $i < j$, $N_{i,j} = \lambda_1 M_1^t diag(\pi_1)X_1N_1 + \lambda_2 M_2^t diag(\pi_2)X_2N_2$, so $N_{1,3} = N_{1,4} = N_{2,3} = N_{2,4}$ and then $N_{x,i}^{-1}N_{x,i^c} = Id$. Therefore $p_k^{12|34} = p = \lambda_1 p_1 + \lambda_2 p_2$ and,

$$Flat_{13|24}\left(p_k^{12|34}\right) = \lambda_1 Flat_{13|24}(p_1) + \lambda_2 Flat_{13|24}(p_2).$$

If the parameters are stochastic then $Flat_{13|24}\left(p_k^{12|34}\right)$ is PSD since it is the sum of two PSD matrices. The equality $Flat_{13|24}\left(p_k^{12|34}\right) = Flat_{14|23}\left(p_k^{12|34}\right)$ follows from Lemma 3.2.9 □

3.3 SAQ: Semi-Algebraic Quartet reconstruction method

In this section we present the phylogenetic quartet reconstruction method SAQ named for *Semi-Algebraic Quartet reconstruction*. It is based on the algebraic and semi-algebraic description of distributions $p \in \mathbb{R}^{4^4}$ that arise from the general Markov model on a quartet T given in Theorem 1.3.24 and Theorem 1.3.26.

Note that Theorem 1.3.24 relies on the tree topology of T however, it does not reflect the stochasticity conditions of the transition matrices attached to T. The stochastic nature of the transition matrix at the interior edge of T is reflected on Theorem 1.3.26.

According to the results in Chapter 2, it is important to combine Theorems 1.3.24 and 1.3.26 in order to obtain successful reconstruction methods for data that might be misleading (that is, small samples or data coming from trees with a short interior edge).

In what follows, we explain how SAQ combines conditions in the theorems mentioned. For each quartet $T_{A|B} \in \mathcal{T}_4$, given a distribution $p \in \mathbb{R}^{4^4}$, SAQ computes a score as follows. If $T = T_{12|34}$, we consider the sixteen 12|34 leaf-transformations $p_k^{12|34}$,

$k = 1, \dots, 16$ of p (see Figure 3.1) and for each of these vectors we compute

$$s^k_{12|34} := \frac{\min\left\{\delta_4\left(\text{psd}\left(Flat_{13|24}\left(p_k^{12|34}\right)\right)\right), \delta_4\left(\text{psd}\left(Flat_{14|23}\left(p_k^{12|34}\right)\right)\right)\right\}}{\delta_4\left(\text{psd}\left(Flat_{12|34}\left(p_k^{12|34}\right)\right)\right) + \delta_4\left(\text{psd}\left(Flat_{12|43}\left(p_k^{12|34}\right)\right)\right)}.$$

Then, define $s_T(p)$ as the average of these sixteen quantities. If T is any of the other two quartets in $\mathcal{T}_4$, $s_T(p)$ is computed analogously by permuting the roles of the leaves accordingly. Specifically,

$$s^k_{13|24} := \frac{\min\left\{\delta_4\left(\text{psd}\left(Flat_{12|34}\left(p_k^{13|24}\right)\right)\right), \delta_4\left(\text{psd}\left(Flat_{14|32}\left(p_k^{13|24}\right)\right)\right)\right\}}{\delta_4\left(\text{psd}\left(Flat_{13|24}\left(p_k^{13|24}\right)\right)\right) + \delta_4\left(\text{psd}\left(Flat_{13|42}\left(p_k^{13|24}\right)\right)\right)},$$

$$s^k_{14|23} := \frac{\min\left\{\delta_4\left(\text{psd}\left(Flat_{12|43}\left(p_k^{14|23}\right)\right)\right), \delta_4\left(\text{psd}\left(Flat_{13|42}\left(p_k^{14|23}\right)\right)\right)\right\}}{\delta_4\left(\text{psd}\left(Flat_{14|23}\left(p_k^{14|23}\right)\right)\right) + \delta_4\left(\text{psd}\left(Flat_{14|32}\left(p_k^{14|23}\right)\right)\right)}.$$

Observe that the denominator of each score s^i_T for $T = T_{ij|kl}$ is computed by doing an average of the flattenings $Flat_{ij|kl}$ and $Flat_{ji|lk}$ since the other permutations of the leaves that preserve the tree topology give rise to redundant flattenings, see Remark 3.2.12. Similarly, the flattenings considered in the numerator of s^k_T are the only non-redundant flattenings that are positive definite of rank 16 for distributions arising on the corresponding tree (see also see Remark 3.2.12).

Finally SAQ outputs the normalized three scores, that is, if $s := s_{12|34}(p) + s_{13|24}(p) + s_{14|23}(p)$, then

$$\texttt{SAQ}(p) := \frac{1}{s}\left(s_{12|34}(p), s_{13|24}(p), s_{14|23}(p)\right).$$

According to SAQ, the correct topology for a distribution p is the topology T for which $s_T(p)$ attains the maximum value. Hereby, SAQ seeks to minimize the average distance δ_4 of the flattening $Flat_{A|B}$ of the $A|B$ leaf-transformations and to maximize the distance δ_4 of the other two flattenings, under the assumption that these should be positive definite. SAQ is a quartet inference measure that satisfies the property I of Definition 1.4.1.

Statistical consistency of SAQ

Let p be a distribution that has arisen on the quartet $T_{12|34}$ with generic stochastic parameters (namely, invertible transition matrices with positive entries and positive distribution at the root node) and let F be the vector of relative frequencies of patterns from N independent trials sampled from p. We know that $\delta_4\left(psd\left(Flat_{12|34}\left(p_k^{12|34}\right)\right)\right)$ and $\delta_4\left(psd\left(Flat_{12|43}\left(p_k^{12|34}\right)\right)\right)$ are zero (by Corollary 3.2.10 (a) and Remark 3.2.12). Moreover, $\delta_4\left(psd\left(Flat_{13|24}\left(p_k^{12|34}\right)\right)\right) > 0$ (analogously for $Flat_{14|23}\left(p_k^{12|34}\right)$) by Corollary 3.2.10 (b). Since $F \to p$ when the length N of the sample tends to infinite, by continuity we infer that $s_{12|34}(F) \to \infty$.

On the other hand, $s_{13|24}(F)$ and $s_{14|23}(F)$ tend to zero when N tends to infinite. Indeed, in order to compute $s_{13|24}(p)$ we consider the sixteen 13|24 leaf-transformations of p; then $Flat_{12|34}\left(p_k^{13|24}\right)$ has still rank ≤ 4 (see Proposition 3.2.7) and its closest positive definite matrix must have rank ≤ 4 by Corollary 3.1.3. Thus,

$$\min\left\{\delta_4\left(psd\left(Flat_{12|34}\left(p_k^{13|24}\right)\right)\right), \delta_4\left(psd\left(Flat_{14|23}\left(p_k^{13|24}\right)\right)\right)\right\} = 0.$$

Moreover, as p arises from generic parameters, $Flat_{13|24}\left(p_k^{13|24}\right)$ has rank 16, is not symmetric anymore and its closest positive definite matrix has generically rank strictly larger than four (we have seen this for the K81 model in Lemma 3.2.11 (b) and we have checked it on simulated random matrices for the GM model). Hence, in this case the distance $\delta_4\left(psd\left(Flat_{13|24}\left(p_k^{13|24}\right)\right)\right)$ is strictly positive. This guarantees that the denominator of $s^k_{13|24}$does not vanish, so $s^k_{13|24}$ is a continuous function and $s_{13|24}(p)$ is zero. A similar argument applies to 14|23.

By normalizing the scores we get that

$$\lim_{N\to\infty} \texttt{SAQ}(F) = \texttt{SAQ}(p) = (1, 0, 0).$$

This immediately gives that the choice of the topology with highest score gives a statistically consistent method.

Theorem 3.3.1. *The topology reconstruction method that assigns to a distribution F the quartet $T_{A|B}$ such that $s_{A|B}(F)$ is maximum is a statistically consistent method for the general Markov model.*

Proof. Let p be a distribution that has arisen on the quartet $T_{12|34}$ with generic stochastic parameters (namely, invertible transition matrices with positive entries and positive distribution at the root node) and let F be the vector of relative frequencies of patterns from N independent trials sampled from p. Following the previous discussion, we have that

$$\lim_{N\to\infty} \mathtt{SAQ}(F) = \lim_{N\to\infty} \left(s_{12|34}(F), s_{13|24}(F), s_{14|23}(F)\right) = \mathtt{SAQ}(p) = (1,0,0).$$

This means that for any $\epsilon > 0$ there exist N_1, N_2, N_3 such that for any $N > N_0 = \max\{N_1, N_2, N_3\}$ we have

$$\left|s_{12|34}(F) - 1\right| < \epsilon, \quad \left|s_{13|24}(F)\right| < \epsilon, \quad \left|s_{14|23}(F)\right| < \epsilon.$$

Thus, if ϵ is smaller than $1/2$ we obtain

$$s_{12|34}(F) > 1 - \epsilon > \epsilon > \max\left\{s_{13|24}(F), s_{14|23}(F)\right\}.$$

Therefore, the probability that the method that chooses the quartet with highest score selects $T_{12|34}$ tends to 1 when N tends to infinite. □

Implementation of the method

In practice, the implementation of an algorithm that computes these scores requires dealing with some technical difficulties. For instance, leaf-transformations might require computing the inverse of some ill-conditioned matrices derived from the marginalization over two taxa. Moreover, when p is an empirical distribution, the corresponding leaf-transformations may not be distributions anymore. The implementation, written in `c++`, that we provide in

`https://github.com/marinagarrote/SAQ-method`

excludes leaf-transformations that are far from being distributions (this is essentially controlled by the parameter *filter* which can be modified by the user and adapted to the alignment length). In Algorithm 3.1 we present a pseudocode for the `SAQ` method.

Algorithm 3.1: SAQ: Semi-Algebraic Quartet reconstruction method

Input: A DNA alignment for 4 taxa of length N and a value for *filter*.

Compute the vector $p \in \mathbb{R}^{4^4}$ of observed relative frequencies of site patterns.;
foreach $T_{A|B} \in \mathcal{T}_4$ **do**
    Set: $\mathcal{S}_{A|B} := \{\}$; // Empty list for the scores $s^k_{A|B}$
    foreach $k = 1, \dots, 16$ **do**
        Apply the k-th $A|B$ leaf-transformations to F to obtain $F^{A|B}_k$;
        if *Every entry of* $F^{A|B}_k >$ *filter* **then**
            Calculate $s^k_{A|B}$ and add it to $\mathcal{S}_{A|B}$;
    if $\mathcal{S}_{A|B}$ *is not empty* **then**
        Calculate $s_{A|B} := \text{mean}\{\mathcal{S}_{A|B}\}$;
    else
        Exit. SAQ can not be used for F;
Calculate $s := s_{12|34}(F) + s_{13|24}(F) + s_{14|23}(F)$;
Calculate $\texttt{SAQ}(F) := \frac{1}{s}\left(s_{12|34}(F), s_{13|24}(F), s_{14|23}(F)\right)$;

Output: Return $\texttt{SAQ}(F)$. SAQ chooses the quartet with maximum $s_{A|B}(F)$.

3.4 THE PARALINEAR METHOD

In this section we introduce the paralinear method that will allow us to use together SAQ and Erik+2 (introduced in Section 1.4.4) in order to take advantage of both methods. To this end we use the paralinear distance and the four-point condition.

Given a distribution vector $p \in \mathbb{R}^{4^4}$ representing the joint distribution of nucleotides patterns of 4 taxa, we compute all pairwise paralinear distances $d_{par}(x, y)$ between the four taxa. That is, consider two taxa x and y (i.e. two coordinates x and y of p) and let J be the underlying joint probability matrix between x and y, this is, every entry (i, j) of J is the probability of observing nucleotides i and j at x and y, respectively (so that the sum of its entries is one). If $\det J \neq 0$, the following value corresponds to the paralinear distance (see Definition 1.1.33) between x and y

$$d_{par}(x, y) = -\log \frac{|\det J|}{\sqrt{\det D_x}\sqrt{\det D_y}}, \tag{3.11}$$

where D_x and D_y are the diagonal matrices whose diagonal entries are given by $J\mathbf{1}$ and

$\mathbf{1}^t J$, respectively. Otherwise, if $\det J = 0$ we take $d_{par}(x,y)$ as infinity. This definition gives a dissimilarity map (see Definition 1.1.34).

The following lemma generalizes the non-negative and the additive properties to general Markov matrices as soon as they are not singular.

Lemma 3.4.1. *Let π be the nucleotide distribution at x and consider a substitution process leading from x to y and ruled by a Markov matrix M. Then, the paralinear distance between x and y defined above (3.11) can be computed as*

$$d_{par}(x,y) = -\log \frac{|\det M| \sqrt{\det\left(diag(\pi)\right)}}{\sqrt{\det\left(diag(M^t\pi)\right)}}. \tag{3.12}$$

Moreover, we have that

(a) $d_{par}(x,y) \geq 0$, and the equality holds if and only if $M = Id$ or is a permutation matrix;

(b) this dissimilarity map $d_{par}(x,y)$ is additive.

Proof. We have that $J = diag(\pi)M$ is the underlying joint probability matrix between x and y so that the sum of its entries is one. Note that $\pi = J\mathbf{1}$ and $\pi^t M = \mathbf{1}^t J$, so $D_x = diag(\pi)$ and $D_y = diag(M^t\pi)$. Therefore,

$$d_{par}(x,y) = -\log \frac{|\det J|}{\sqrt{\det D_x}\sqrt{\det D_y}} = -\log \frac{|\det M|\ \det D_x}{\sqrt{\det D_x}\sqrt{\det D_y}} =$$
$$= -\log \frac{|\det M| \sqrt{\det\left(diag(\pi)\right)}}{\sqrt{\det\left(diag(M^t\pi)\right)}}.$$

Now, we proceed to prove *a*): $d_{par}(x,y) \geq 0$ or equivalently, that

$$(\det M)^2 \ \frac{\det diag(\pi)}{\det diag(M^t\pi)} \leq 1. \tag{3.13}$$

First of all, since the matrix M is positive, we deduce that

$$|\det M| = \left| \sum_{\{i_1,i_2,i_3,i_4\}=[4]} \prod_j sgn(i_1,i_2,i_3,i_4)\, m_{i_j,j} \right| \leq \sum_{\{i_1,i_2,i_3,i_4\}=[4]} \prod_j m_{i_j,j}.$$

Then, we have that

$$\begin{aligned}|\det M|\det diag(\pi) &= |\det M| \prod_k \pi_k \leq \prod_k \pi_k \left(\sum_{\{i_1,i_2,i_3,i_4\}=[4]} \prod_j m_{i_j,j} \right) \\ &= \sum_{\{i_1,i_2,i_3,i_4\}=[4]} \prod_j m_{i_j,j}\, \pi_j \leq \sum_{i_1,i_2,i_3,i_4\in[4]} \prod_j m_{i_j,j}\, \pi_{i_j} \\ &= \prod_j \sum_i m_{ji}\pi_j = \prod_i (M^t\,\pi)_i = \det diag(M^t\pi).\end{aligned}$$

By multiplying this inequality with $|\det M| \leq 1$ (that follows since M is a stochastic matrix, see Theorem 1.2.15), we obtain (3.13). Note that if $d_{par}(x,y) = 0$, then necessarily $\det M = 1$. The Perron-Frobenious theory (see Chapter 8 in Meyer, 2000) implies that a stochastic matrix M with determinant 1 is necessarily the identity matrix or a permutation matrix.

Finally, the statement (b) follows easily from the expression (3.12). □

Then, given a distribution $p \in \mathbb{R}^{4^4}$ and a split $A|B$ of the set $[4]$, $A = \{a,b\}, B = \{c,d\}$, we define the following quantity

$$\begin{aligned}I_{A|B}(p) = \min\{&d_{par}(a,c) + d_{par}(b,d), d_{par}(a,d) + d_{par}(b,c)\} \\ &- d_{par}(a,b) - d_{par}(c,d).\end{aligned} \tag{3.14}$$

The quantity above is the "neighborliness" measure used in Gascuel (1994), the "paralinear method" used in Lake (1994), and was presented by Buneman (1971) as a measure of twice the length at the interior edge.

Remark 3.4.2. Keeping the Notation 3.2.1, write $N_{x,y}$ for the matrix consisting on the double marginalization of p on the coordinates different from x and y, with rows labeled by the states at x if $x < y$, or the transpose of this one if $y > x$. Then the quantity $I_{A|B}(p)$ can also be defined as

$$I_{A|B}(p) = -\log \max\left\{ \left| \frac{\det(N_{ac})\det(N_{bd})}{\det(N_{ab})\det(N_{cd})} \right|, \left| \frac{\det(N_{ad})\det(N_{bc})}{\det(N_{ab})\det(N_{cd})} \right| \right\} \tag{3.15}$$

where $A = \{a,b\}$, $B = \{c,d\}$. This follows easily since $J = N_{x,y}$. Therefore, $d_{par}(x,y) = -\log \frac{det(N_{x,y})}{\sqrt{det(N_x)det(N_y)}}$ where $N_x = diag(N_{x,y}\mathbf{1})$ and $N_y = (\mathbf{1}^t N_{x,y})$. This alternative presentation of the definition of $I_{A|B}(p)$ is closer to our original approach and our implementation of the method is based on it.

It is worth pointing out that at most one of the three values $I_{12|34}(p)$, $I_{13|24}(p)$, and $I_{14|23}(p)$ is strictly positive. This is an immediate consequence of (3.14), see also Lemma 8 of Buneman (1971).

Lemma 3.4.3. *Assume that the determinant of every double marginalization of a distribution $p \in \mathbb{R}^{4^4}$ is non-zero (so that all values $d_{par}(x,y)$, with $x,y \in [4]$ can be computed). Then, $I_{A|B}(p) > 0$ for at most one bipartion $A|B$.*

Remark 3.4.4. The value $I_{A|B}$ can be seen as a Markov invariant, see Definition 1.2.20. Indeed, it is straightforward to see that, for any distribution p, $I_{A|B}(p)$ is invariant by the Markov action: $I_{A|B}((N_1 \otimes \cdots \otimes N_4)\cdot p) = I_{A|B}(p)$ for any matrices $N_i \in \mathbb{M}_4^*$.

Note that, in the case that p has arisen on the quartet $T_{A|B}$ with certain stochastic parameters then,

$$I_{A|B}(p) = -\log\det(M_5)^2 \left| \frac{\det(diag(\pi))}{\det(diag(M_5^t\pi))} \right|$$

where M_5 is the transition matrix at the interior edge (see Remark 3.4.2). Note $I_{A|B}(p)$ is non-negative and that under the assumption that M_5 is DLC, we deduce that if $I_{A|B}(p) = 0$, then M_5 is necessarily the identity matrix (see (a) of Lemma 3.4.1).

We denote by $\mathcal{I}(p)$ the triplet of the values $I_{A|B}(p)$:

$$\mathcal{I}(p) = (I_{12|34}(p), I_{13|24}(p), I_{14|23}(p)).$$

When p arises from a quartet tree $T_{A|B}$ under the GM model for some positive substitution parameters we can generalize Lake's results and prove that $d_{par}(x,y) \geq 0$ for any two nodes x, y in the tree, see Theorem 3.4.5 (Lake proved this only for incremental Markov matrices). Moreover, if the entries of the Markov matrix at the interior edge are strictly positive, from the four-point condition (see Theorem 1.1.35) we get that both quantities inside the minimum of (3.14) are equal, $I_{A|B}(p)$ is the unique positive quantity in the triplet and its value coincides with twice the paralinear distance of the interior edge (see Theorem 3.4.5).

Statistical consistency of the paralinear method

In the next Theorem 3.4.5 we see that $\mathcal{I}$ is a quartet-inference measure that satisfies the "strong property II" introduced in Section 1.4. Based on this, we consider the *paralinear method*: given the vector F of relative frequencies of patterns in an alignment, choose the tree $T = T_{A|B}$ with largest value $I_{A|B}(F)$.

Theorem 3.4.5. *Let F be the vector of relative frequencies of patterns from N independent trials sampled from a distribution p, $F \sim Mult(N; p)$. Then,*

$$\mathcal{J}(F) = (I_{12|34}(F), I_{13|24}(F), I_{14|23}(F))$$

is a quartet inference measure that satisfies the property I and the strong property II with $\lambda = 0$ (see Definition 1.4.1).

(a) the limit of the expectation of $\mathcal{J}(F)$ as N goes to infinity is

$$\lim_{N\to\infty} \mathbb{E}(\mathcal{J}(F)) = \big(I_{12|34}(p), I_{13|24}(p), I_{14|23}(p)\big);$$

(b) suppose p has arisen on $T = T_{12|34}$ with stochastic parameters. If u and v are the two interior nodes of T, then $I_{12|34}(p) = 2d_{par}(u, v) \geq 0$ and $I_{C|D}(p) = -I_{12|34}(p)$ for any other split $C|D \neq 12|34$;

(c) the topology reconstruction method that associates to a frequency vector F the tree T with largest $I_T(F)$ is statistically consistent.

Proof. That $\mathcal{J}(F) = (I_{12|34}(F), I_{13|24}(F), I_{14|23}(F))$ is a quartet inference measure that satisfies the property I is direct by definition. The strong property II is satisfied with $\lambda = 0$ because of Remark 3.4.4.

(a) This follows since $\lim_{N\to\infty} \mathcal{J}(F) = \mathcal{J}(p)$ and by Taylor expansion we have

$$\lim_{N\to\infty} \mathbb{E}(\mathcal{J}(F)) = \mathcal{J}(p) = \big(I_{12|34}(p), I_{13|24}(p), I_{14|23}(p)\big);$$

(b) The first claim is a consequence of the additivity of the paralinear distance (b) in Lemma 3.4.1. Indeed, under the assumption that p arises from the tree $T = 12|34$, every quantity $d_{i,j}$ $(i, j \in [4])$ can be written as the sum of the (paralinear) distances attached to the edges of T between i and j. For example, $d_{1,2} = d_{1,u} + d_{u,v} + d_{v,3}$. Then because of the 4-point condition, the value $I_{A|B}(p)$ results in $2d_{u,v}$, which is non-negative in virtue of (a) of Lemma 3.4.1. Now, if $C|D$ is a split other that $12|34$, it is immediate from the Definition (3.14) that $I_{C|D}(p) = -I_{12|34}(p)$.

(c) Denote $\delta = 2d_{u,v}$. Then, following (a) we have that

$$\lim_{N\to\infty} \mathcal{J}(F) = \lim_{N\to\infty} \big(I_{12|34}(F), I_{13|24}(F), I_{14|23}(F)\big) = \mathcal{J}(p) = (\delta, -\delta, -\delta),$$

where the last equality follows from (b). Therefore, for any $\epsilon > 0$ there exist N_1, N_2, N_3 such that for any $N > N_0 = \max\{N_1, N_2, N_3\}$ we have

$$\left|I_{12|34}(F) - \delta\right| < \epsilon, \quad \left|I_{13|24}(F) + \delta\right| < \epsilon, \quad \left|I_{14|23}(F) + \delta\right| < \epsilon.$$

Then, for any $\epsilon < \delta$ we have

$$I_{12|34}(F) > \delta - \epsilon > 0 > -\delta + \epsilon > \max\left\{I_{13|24}(F), I_{14|23}(F)\right\}.$$

Then we can conclude that the method that chooses the tree T with largest $I_T(F)$ chooses $T_{12|34}$ with probability tending to 1 when N tends to infinite. □

We have seen that as a topology reconstruction method, the paralinear method is statistically consistent for the general Markov model (see Theorem 3.4.5 (c)). Moreover, it is highly successful, as indicated by the results on simulated data displayed in Figure 4.11 (see the Section 4.3 section for details). Nevertheless, as we will see in the next chapter, we found that using $\mathcal{I}$ in order to ratify or not the results of `Erik+2` and combine them with `SAQ` gives a better performance on topology reconstruction. How to combine `SAQ`, `Erik+2` and the paralinear method will be done in the next section.

3.5 ASAQ: Algebraic and Semi-Algebraic Quartet reconstruction method

In our simulation studies on the performance of `SAQ` versus phylogenetic reconstruction (see Chapter 4) we note that while `SAQ` usually outperforms `Erik+2` for short alignments (length ≤ 1000), `Erik+2` obtains better results as the length of the alignment increases. This consideration leads us to introduce `ASAQ`, a new combined method of `Erik+2` and `SAQ` that tries to apply one or the other according to whether the data vector is consistent with the positivity of the estimated interior branch length. Therefore, we propose the following quartet reconstruction method `ASAQ` from *Algebraic and Semi-Algebraic Quartet reconstruction method*. Given a distribution $F \in \mathbb{R}^{4^4}$ the new method `ASAQ` applies the `Erik+2` and the paralinear methods to F and checks whether both methods output the same quartet and

(1) if `Erik+2` and $\mathcal{I}$ agree, then `ASAQ` outputs the topology and weights of `Erik+2`,

(2) if they do not agree, then `ASAQ` applies the `SAQ` method and outputs the weights of `SAQ`.

Note that, an inconsistency between Erik+2 and $\mathcal{J}$ implies that the algebraic conditions used by Erik+2 are not in concordance with the fact that the substitution parameters must be *stochastic* (note the role of the positiveness of the entries of the Markov matrix to prove that $d_{par}(x,y) \geq 0$ in (a) of Lemma 3.4.1). In this case ASAQ relies on SAQ which, unlike the method Erik+2, takes into account both the semi-algebraic and the algebraic constraints. ASAQ is a quartet inference measure that satisfies Property I of Definition 1.4.1 and it is as well, a statistically consistent reconstruction method for the general Markov model:

Theorem 3.5.1. *ASAQ is statistically consistent for the general Markov model.*

Proof. The statistical consistency of ASAQ follows immediately from the statistical consistency of Erik+2 (Fernández-Sánchez and Casanellas, 2016) and the paralinear method (see Theorem 3.4.5). That is, if F is the vector of relative frequencies of patterns from N independent trials sampled from a distribution p that has arisen from $T = T_{12|34}$ with non-singular parameters, $F \sim \text{Mult}(N; p)$, then the probability that both methods, when applied to F, select $T_{12|34}$ tends to 1 when N tends to infinite. □

One can use ASAQ with mixtures of distributions on the same tree topology with two or three categories (or partitions) as this is allowed in Erik+2 (see Section 1.4.4). However, we cannot claim that ASAQ is statistically consistent for mixtures because $\mathcal{J}$ and SAQ are not consistent in this scenario (a priori). However, the simulation studies in Section 4.2.3 show a good performance of SAQ on mixture data obtained from the same tree, and this leads to a good performance of ASAQ as well (see Section 4.4.3). The limit on the number of categories ($m = 3$) comes from the theoretical foundations of Erik+2, as a larger amount of categories would make impossible the identifiability of the tree topology in terms of the rank constraints of the flattenings.

Implementation of the method

Given a distribution $F \in \mathbb{R}^{4^4}$ obtained from an alignment and keeping the notation already used, the computation of $I_{A|B}(F)$ where $A = \{a, b\}$ and $B = \{c, d\}$ is done as:

$$\begin{aligned} I_{A|B}(p) = \min\{&-\log|\det(N_{ac})| - \log|\det(N_{bd})|, -\log|\det(N_{ac})| - \log|\det(N_{bd})|\} \\ &+ \log|\det(N_{ab})| + \log|\det(N_{cd})|. \end{aligned} \tag{3.16}$$

In order to avoid numerical problems, we compute $I_{A|B}(p)$ only if the condition number of the matrices $N_{x,y}$ involved are less than a certain prescribed tolerance. To this end,

we use the parameter "threshold" which is set to 5000, but can be modified by the user, for example to adapt it to the alignment length. If $I_{A|B}(p)$ cannot be computed, this may be an indicator of short sample size; in this case, ASAQ outputs the topology and the weighting system given by SAQ.

We present a pseudocode for the ASAQ method in Algorithm 3.2. The code is written in C++ and can be downloaded from

https://github.com/marinagarrote/ASAQ-method.

Algorithm 3.2: ASAQ: Algebraic and Semi-Algebraic Quartet reconstruction method

Input: A DNA alignment for 4 taxa of length N and a value for *filter* and a *threshold*.

Compute the vector $F \in \mathbb{R}^{4^4}$ of observed relative frequencies of site patterns.;
Calculate the Erik+2 weights:
$\texttt{Erik+2}(F) = \frac{1}{w}(w_{12|34}(F), w_{12|34}(F), w_{12|34}(F))$; `// see Section 1.4.4`
Select the tree $T_{A|B}$ that maximizes Erik+2(F);
Compute the marginalizations $N_{i,j}$;
if $condNumb(N_{i,j}) < threshold$ **then**
    Calculate $I_{A|B}(F)$ as in (3.16);
if $\exists I_{A|B}(F)$ & $I_{A|B}(F) > 0$ **then**
    Return $T_{A|B}$ and Erik+2(F);
else
    Calculate $\texttt{SAQ}(F) := \frac{1}{s}\left(s_{12|34}(F), s_{13|24}(F), s_{14|23}(F)\right)$;
    Return $T_{A|B}$ that maximizes SAQ(F) and the weights SAQ(F);

4

RESULTS ON REAL AND SIMULATED DATA

We have tested the two methods SAQ and ASAQ introduced in the previous chapter under different settings of simulated data: on a "tree space" of quartets, on quartets with random branch lengths, and on alignments that are not identically distributed across sites. We also considered the three quartet-based methods (Q-methods) described in Section 1.4.5: Quartet Puzzling (Strimmer and Haeseler, 1996), Willson's method (Willson, 1999) and Weight Optimization (Ranwez and Gascuel, 2001) and analysed their performance with the weighting system of ASAQ introduced in the previous chapter. In this chapter we present the performance of all these methods under different data sets (including real data) and compare it to other phylogenetic reconstruction methods such as maximum likelihood (ML), neighbor-joining (NJ), maximum parsimony (MP) and Erik+2 (see Sections 1.4.1, 1.4.2, 1.4.3 and 1.4.4).

The organization of this chapter is as follows. First we describe the different data sets that will be used to analyse the performance of our methods. In Section 4.2 we present the performance of SAQ as a quartet reconstruction method on different types of simulated and real data and compare it with the performance of other inference methods. Next, we present the performance of the paralinear method on the "tree space". Section 4.4 is devoted to present the performance of ASAQ as a quartet inference method and in Section 4.5 we present the performance of the Q-methods mentioned using weighting systems provided by ML, NJ, Erik+2 and ASAQ. Finally we discuss the results and main conclusions regarding methods SAQ and ASAQ.

Computations

The computations on this chapter have been performed on a computer with 6 Dual Core Intel(R) Xeon(R) E5-2430 Processor (2.20 GHz) equipped with 25 GB RAM running Debian GNU/Linux 8. We have used the g++ (Debian 4.9.2-10+deb8u2) 4.9.2 compiler and the C++ library for linear algebra & scientific computing *Armadillo* version 3.2.3 (Creamfields).

The average time needed to apply `SAQ` and `ASAQ` to 100 alignments of length 10 000 bp is 8.5 and 8.7 seconds, respectively. For the same data, the paralinear method takes 7.8 seconds. The average time to reconstruct one CD tree given weights for its 495 quartets subtrees is 1.8 seconds for QP, 89.5 seconds for WIL and 4 seconds for WO.

4.1 DATA

We considered two different scenarios of simulated data: one for testing `SAQ` and `ASAQ` as quartet reconstruction methods (described in Section 4.1.1) and another for testing the Q-methods with several weighting systems (see Section 4.1.2). For the first one we use the simulated data described by Fernández-Sánchez and Casanellas (2016) whereas for the second we follow the approach of Ranwez and Gascuel (2001) and consider 12-leaf trees.

Simulated data has been generated according to either a general Markov model (see Definition 1.1.30) or a homogeneous general time-reversible model (see Definition 1.1.31):

- To simulate data under the GM model, we have used the software `GenNon-h` (Kedzierska and Casanellas, 2012) that given a set of branch lengths and a quartet T, generates a random distribution of nucleotides at the root and random substitution matrices with the expected branch length (approximated as in (1.12)). Then it produces multiple DNA sequence alignments that have evolved under the Markov processes on T with the corresponding substitution parameters.
- When we refer to GTR we mean a homogeneous GTR, i.e. the same rate matrix is assumed at all edges of the tree. Data evolving under the homogeneous GTR model have been generated using `Seq-gen` (Rambaut and Grass, 1997).

4.1.1 Simulated data for quartet reconstruction

In this section we describe different type of quartets and parameters and the data generated according to them.

Tree space

The first data set we use corresponds to the tree space suggested by Huelsenbeck (1995). We consider quartets as Figure 4.1 (a), whose branch lengths are given by a pair of parameters a and b which vary between 0 and 1.5 in steps of 0.02. The resulting *tree space* is shown in the Figure 4.1 (b). The upper left region of this tree space corresponds to the "Felsenstein zone", which contains trees subject to the long branch attraction phenomenon, see Section 1.4.3.

For each of the two nucleotide substitution models considered, GM and GTR, and for each pair (a, b), we have simulated one hundred alignments with the corresponding branch lengths. The alignment lengths considered are 500, 1 000 and 10 000 sites or base pairs (bp for short).

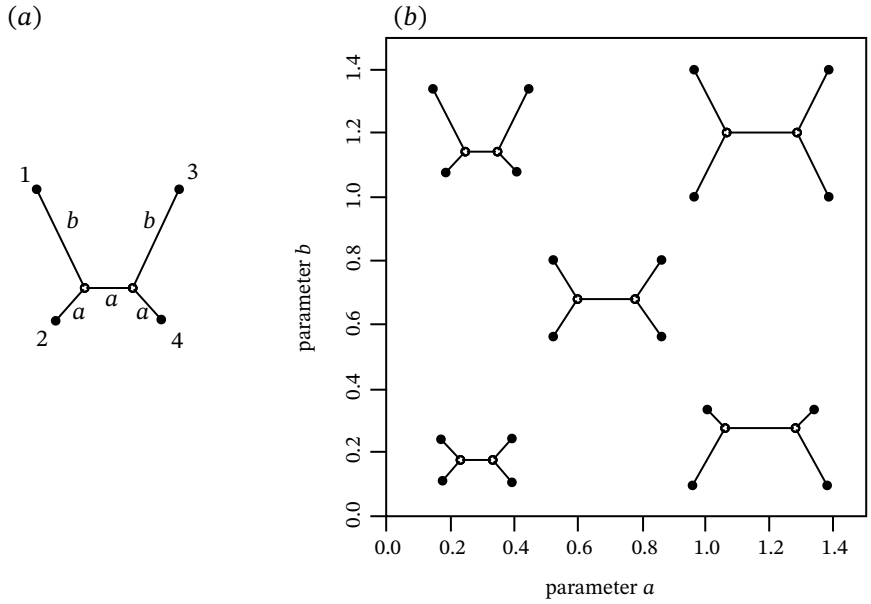


FIGURE 4.1: (a) 4-leaf tree where the length of two opposite branches and the interior branch are represented by a; the other two peripheral branches have length b. Branch lengths are measured in the expected number of substitutions per site . (b) Tree space obtained from the tree in (a) when the branch lengths a and b are varied from 0.01 to 1.5 in steps of 0.02.

Random branch lengths

The second data set consists on 10 000 alignments generated from quartets whose branch lengths are randomly generated according to a uniform distribution in the intervals $(0, 1)$ or $(0, 3)$. These alignments are obtained according to both substitution models, GM and GTR, and are either 1 000 or 10 000 bp long.

Mixture data

We also test our methods on data from a mixture of distributions, even if they are not specifically designed to deal with this type of mixtures. We consider a mixture alignment with two categories of the same size, both evolving under the GM model on the same quartet topology: in the notation of Figure 4.1 (a), the first category corresponds to external branch lengths $a = 0.05$, $b = 0.75$, while the second corresponds to $a = 0.75$ and $b = 0.05$. The internal branch length, parametrized by r, takes the same value in both categories and varies from 0.01 to 0.4 in steps of 0.05 (see Figure 4.2). The total length of the alignments considered is 1 000 and 10 000 bp. This approach was first used by Kolaczkowski and Thornton (2004) to discuss the performance of `ML` and `MP` when data violate the assumption of the method.

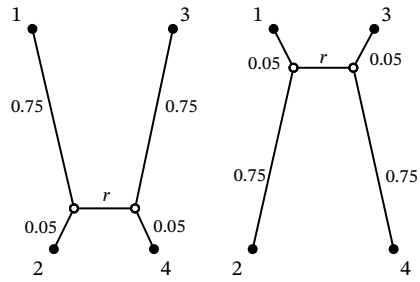


FIGURE 4.2: Trees considered in the categories of the alignment in the simulations on mixture data: the alignment has two categories of the same size, each evolving under the GM model on one of the trees depicted above with the systems of branch lengths indicated. The internal branch length takes the same value in both categories is varied from 0.01 to 0.4 in steps of 0.05.

4.1.2 Simulated data for larger trees

We followed Ranwez and Gascuel (2001) to test the performance of different Q-methods with different weighting systems. To this end, we considered three tree topologies on 12-leaf trees, denoted *CC*, *CD* and *DD*, with some prescribed proportion among their edge lengths (see Figure 4.3). A parameter b controls the branch lengths and takes values 0.005, 0.015, 0.05, and 0.1, producing a maximum pairwise divergence along the tree of about 0.125, 0.375, 1.25, and 2.5 substitutions per site, respectively.

We simulated data under the GM model on the three topologies. However, we only considered data evolving under GTR model in the case of the topology *DD* since it is arguably the hardest to reconstruct, see the Section 4.4 and Ranwez and Gascuel (2001).

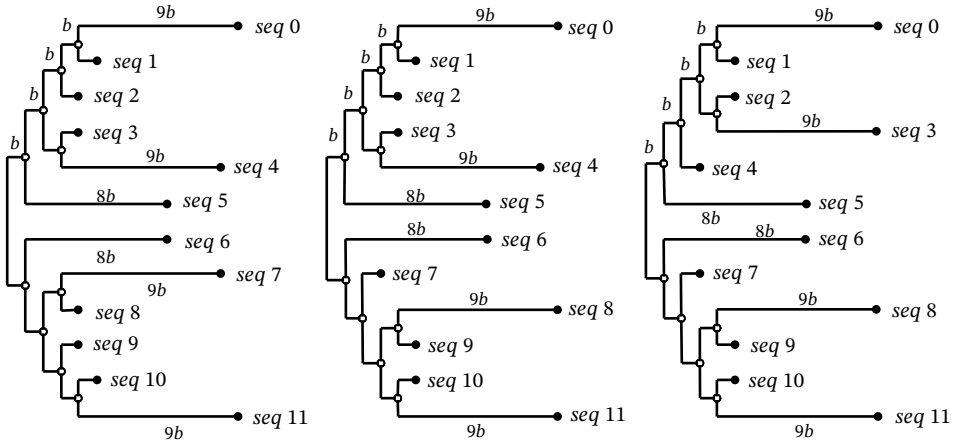


FIGURE 4.3: The three different tree topologies *CC*, *CD* and *DD* on 12 taxa considered to test the Q-methods with different weighting systems. They are obtained by glueing a combintation of two trees (*C* and *D*) by the root. Here, the parameter *b* represents the length of internal branches.

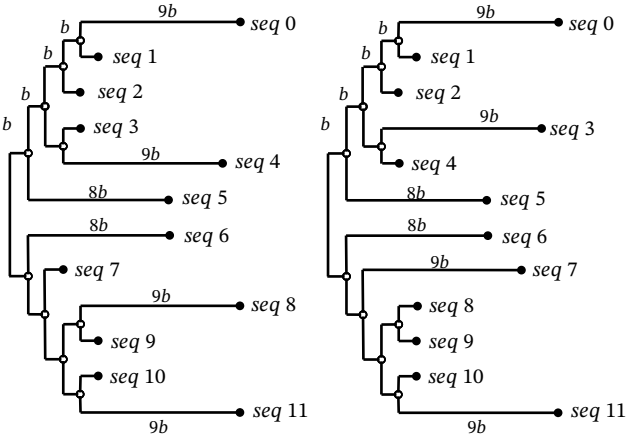


FIGURE 4.4: Trees considered in the categories of the alignment in the simulations on mixture data for Q-methods: the alignment has two categories and $p \in \{0.25, 0.50, 0.75\}$ represents the proportion of the first category relative to the second. Data are generated under the GM model on one of the trees depicted above with the systems of branch lengths indicated for different values of *b*.

We also considered a 2-category mixture model on these trees as follows. We generated alignments on the tree topology *CD* evolving under a GM model but with two systems of substitution parameters. The sites of the first category, namely a fraction of *p* sites, were generated assuming the branch lengths of the *CD* tree on the left of Figure 4.4 (which depend on the internal length *b* as explained above). For the rest of the sites,

we exchange the branch lengths of the edges leading to $seq3$, $seq4$ and $seq7$, $seq8$ respectively as presented on the right of Figure 4.4. We varied the proportion p of sites in the first category between 0.25, 0.50, and 0.75. The length of the interior branch length b is varied as explained above.

We have considered 100 alignments of lengths 600 (in order to match the alignment length considered in Ranwez and Gascuel, 2001), 5 000 and 10 000 base pairs generated according to GM or GTR.

4.1.3 Real data

We considered data of 106 orthologous genes of eight species of yeast, *Saccharomyces cerevisiae*, *S. paradoxus*, *S. mikatae*, *S. kudriavzevii*, *S. castellii*, *S. kluyveri*, *S. bayanus*, and *Candida albicans*. The alignment consisted on the concatenation of 42 337 second codon positions of these genes as provided by Jayaswal et al. (2014). These data sets were originally studied by Rokas et al. (2003) who identified the tree topology T of Figure 4.5 (left) with 100% bootstrap support. This tree is widely accepted by the community of biologists but its correct inference is known to depend on the consideration of heterogeneity across lineages (Rokas et al., 2003; Phillips, Delsuc, and Penny, 2004; Jayaswal et al., 2014). An inaccurate underlaying model usually leads to the reconstruction of the tree topology T' of Phillips, Delsuc, and Penny (2004) (see Figure 4.5 right).

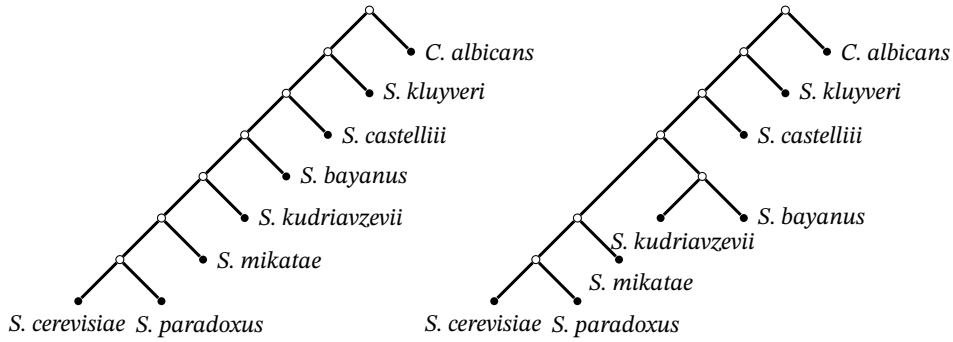


FIGURE 4.5: The tree T of Rokas et al. (2003) (left) and the alternative tree T' of Phillips, Delsuc, and Penny (2004) (right) are constructed using the data provided by Jayaswal et al. (2014) with 42 337 second codon positions of 106 orthologous genes of *Saccharomyces cerevisiae*, *S. paradoxus*, *S. mikatae*, *S. kudriavzevii*, *S. castellii*, *S. kluyveri*, *S. bayanus*, and *Candida albicans*.

4.2 PERFORMANCE OF SAQ ON QUARTETS

In this section we present the performance of SAQ on the different scenarios of quartet data presented in the above section.

4.2.1 SAQ on the tree space

The performance of SAQ on the tree space data for alignments of 500, 1 000 and 10 000 bp simulated under GM or GTR data is presented in Figure 4.6. The success of SAQ in recovering the correct quartet is represented by different tones of gray, where black corresponds to 100% success and white corresponds to 0% success. Gray tones correspond to regions of intermediate probability, and the 95 % and 33 % isoclines are represented with a white and black line respectively.

We observe that these figures exhibit a consistent performance (according to the results in Huelsenbeck, 1995 and Fernández-Sánchez and Casanellas, 2016), with a decreasing performance at the Felsenstein zone but a high performance at the other zones, and with an increase of success for larger samples.

Average success of different quartet methods on the tree space of Figure 4.1 (b)

simulations	base pairs	SAQ	Erik+2	NJ	ML
GM	500	**84.6**	72.4	72.5	72.1
	1 000	**88.8**	80.3	79.7	73.6
	10 000	96.8	**97.1**	94.3	75.4
GTR	500	78.4	74.8	72.9	**88.0**
	1 000	83.5	84.3	80.5	**93.4**
	10 000	94.5	**99.2**	94.5	98

TABLE 4.1: Average success of SAQ corresponding to the tree space of Figure 4.1 (b), compared with the results of the performance of Erik+2, neighbor-joining (NJ) and maximum likelihood (ML) taken from Fernández-Sánchez and Casanellas (2016). ML(homGMc) estimates a homogeneous continuous GM model (that is, it estimates an unrestricted rate matrix for the whole tree and a distribution at the root) and is applied when data is generated under a GM model, while ML(GTR) is applied when data are generated under GTR. In each row of the table, the highest success is indicated in bold font.

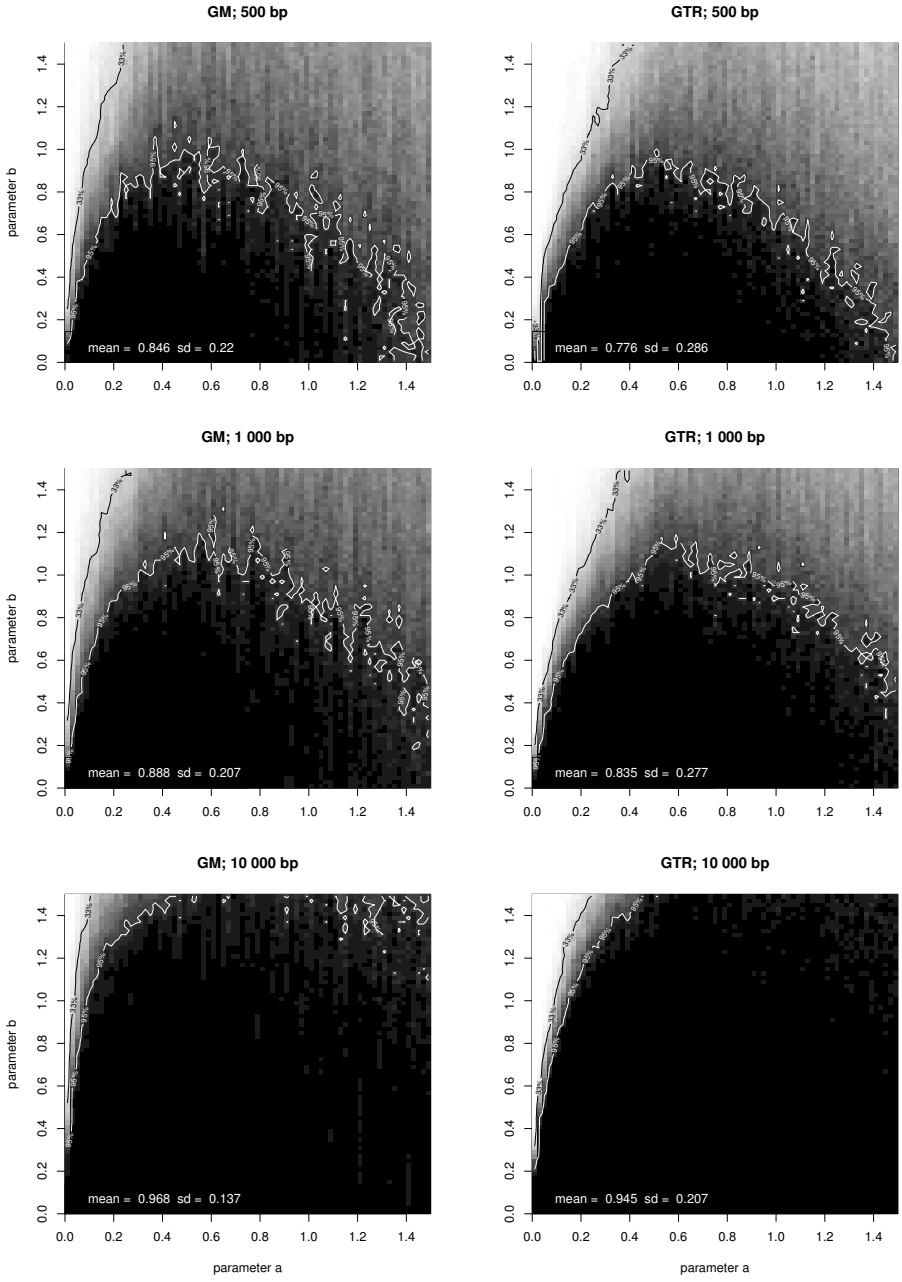


FIGURE 4.6: Performance of SAQ in the tree space of Figure 4.1 (b) on alignments of length 500 bp (top), 1 000 bp (middle) and 10 000 bp (bottom). Black is used to represent 100% of successful quartet reconstruction, white to represent 0%, and different tones of gray the intermediate frequencies. The 95% contour line is drawn in white, whereas the 33% contour line is drawn in black. Left: data generated under the GM model; Right: data generated under a homogenous GTR model.

In Table 4.1 we summarize the overall performance of SAQ on the tree space in comparison with the methods studied in Fernández-Sánchez and Casanellas (2016). In particular, we compare it with a maximum likelihood approach (see Section 1.4.1): ML(homGMc) estimates the most general homogeneous across lineages continuous-time model and ML(GTR) estimates a homogeneous across lineages GTR model. Both methods estimate the rate matrix and the distribution at the root. We also compare it to the neighbor-joining method (NJ) under the paralinear distance, as was used in the quoted paper. We also provide the comparison to the algebraic method Erik+2 explained in Section 1.4.4. The plots of the performance of these methods on the tree space described above can be found in Fernández-Sánchez and Casanellas (2016).

For simpler substitution models, such as JC69, K80 and K81, the rank conditions can be refined to obtain other algebraic conditions that take into account the particular symmetries of these models, see Casanellas and Fernández-Sánchez (2011). It would be feasible to obtain different versions of SAQ (and ASAQ) tailored specifically to these models; however, in the absence of such tailoring, other methods (such as ML) can take advantage of being model-specific and outperform the current version of SAQ for data consistent with some particular surrogate model. We have tested SAQ on the tree space of Figure 4.1 with data simulated from the heterogeneous K81 model across lineages (generated with GenNon-h by Kedzierska and Casanellas, 2012) and obtained an overall performance of 83.7% for 1 000 bp and 93.6% for 10 000 bp, allowing us to conclude that performance drops only slightly when moving from GM to K81.

4.2.2 SAQ on trees with random branch lengths

As the weights of SAQ are normalized so that the three quartet values sum to one we can represent these scores in a ternary plot (also called a simplex plot), see Strimmer and Haeseler (1997). Every point in a ternary plot corresponds to a triplet of weights of SAQ. For instance, the vertex at bottom left corresponds to the triplet $(1, 0, 0)$ while the centroid of the triangle is $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$.

To visualize how the output of SAQ is distributed, Figures 4.7 and 4.8 show ternary plots corresponding to the data described in Section 4.1.1.

We observe that SAQ weights are equally distributed for the two wrong topologies and the vast majority of points lie close to the left corner (which represents the correct quartet) for both GM and GTR data. We also observe a strong difference in the performance of SAQ when applied to branch lengths in (0,1) or in (0,3), being notably higher in the former. The average success of SAQ for these random branch systems is shown in Table 4.2.

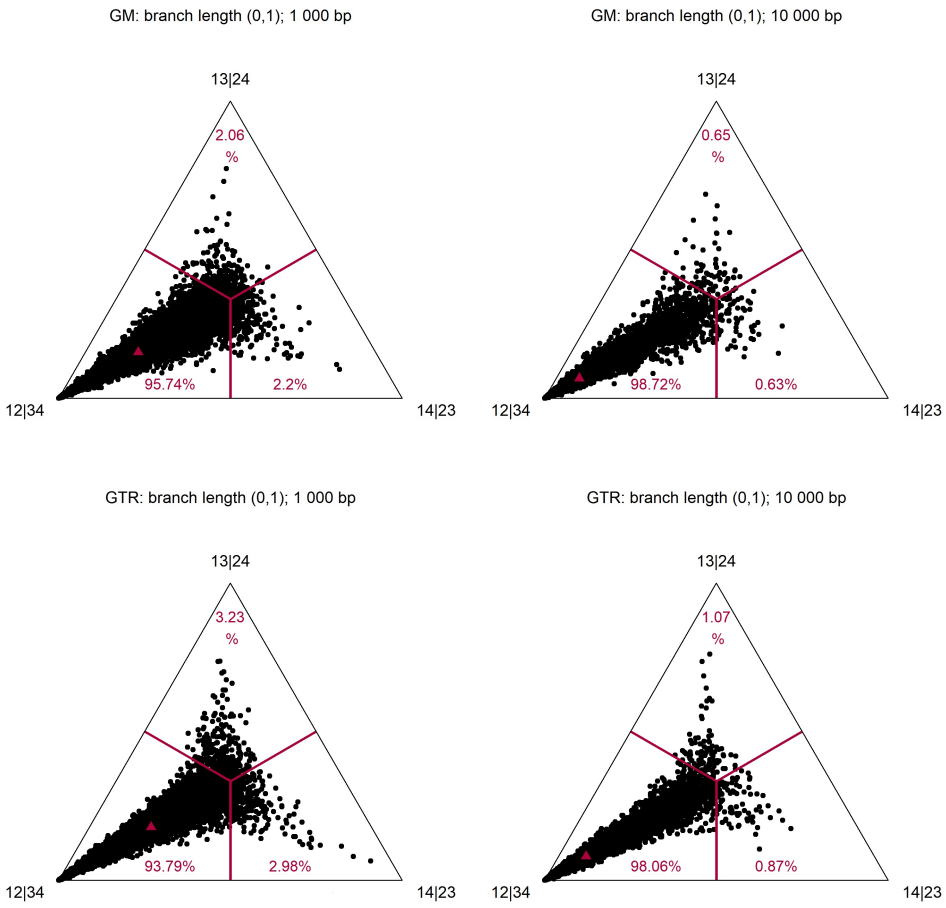


FIGURE 4.7: Ternary plots corresponding to the weights of `SAQ` applied to 10 000 alignments generated under the 12|34 tree with random branch lengths uniformly distributed between 0 and 1. On each triangle the bottom-left vertex represents the underlying quartet 12|34, the bottom-right vertex is the quartet 13|24 and the top vertex is 14|23. Each triangle is divided into three regions according to which quartet is selected by the method and the percentage of trees indicated on that region is shown in red. The red triangle corresponds to the point obtained by computing the average of `SAQ` weights applied to the 10 000 alignments. Top: data generated under GMM; bottom: data generated under GTR. Left : 1 000 bp; Right: 10 000 bp.

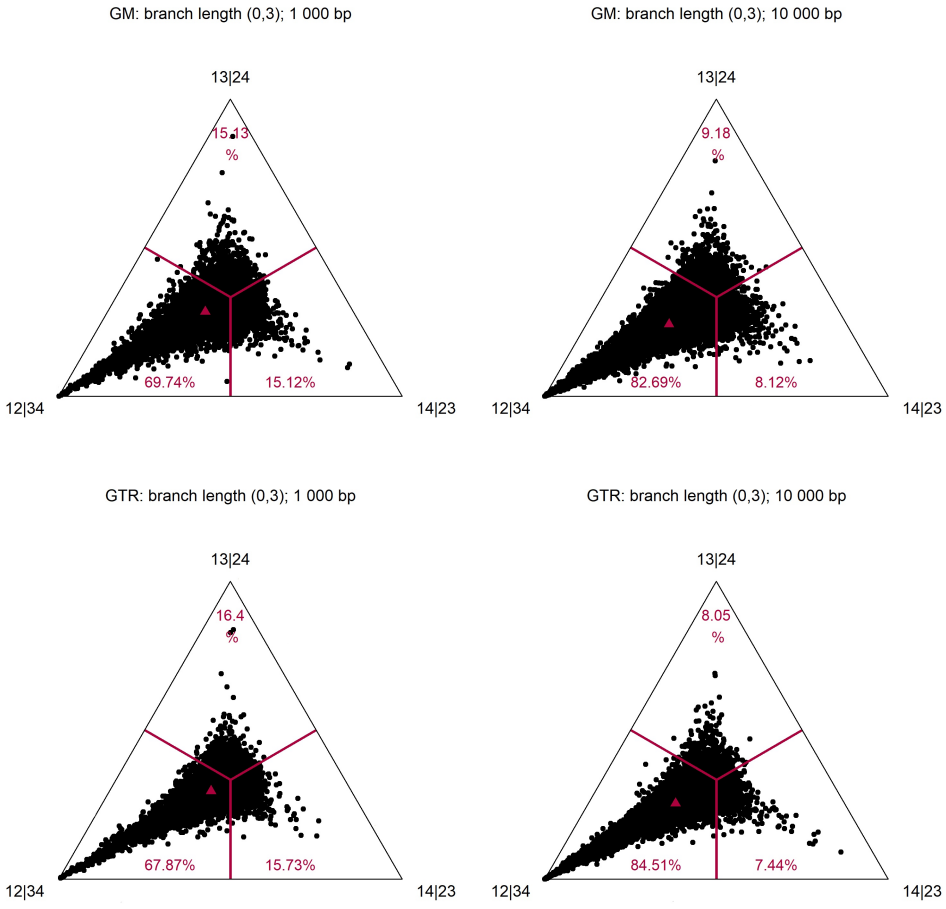


FIGURE 4.8: Ternary plots corresponding to the weights of SAQ applied to 10 000 alignments generated under the 12|34 quartet with random branch lengths uniformly distributed between 0 and 3. On each triangle the bottom-left vertex represents the underlying true topology 12|34, the bottom-right vertex is topology 13|24 and the top vertex is 14|23. See caption of Figure 4.7 for a complete description. Top: correspond to data generated under GM model; bottom: data generated under homogenenous GTR. Left : 1 000 bp; Right: 10 000 bp.).

Average success of SAQ applied to data generated on 12|34 with random branch lengths

branch length	GM		GTR	
	1 000 bp	10 000 bp	1 000 bp	10 000 bp
(0,1)	95.74	98.72	93.72	98.06
(0,3)	69.74	82.69	67.87	84.51

TABLE 4.2: Average success of SAQ on alignments of lengths 1 000 and 10 000 bp generated on the quartet 12|34 under the GM and GTR models with random branch lenghts uniformely distributed in (0,1) (first row) and (0,3) (second row).

4.2.3 SAQ on mixture data

The performance of the method SAQ under the mixture data described in Section 4.1.1 is shown in Figure 4.9. The ternary diagrams show a high accuracy in determining the correct quartet, even when the length of the alignments is 1 000 bp.

In the same figure, we present the performance of SAQ in terms of the branch length at the interior edge, following the study in Kolaczkowski and Thornton (2004) for lengths 1 000, 10 000 and 100 000 bp. This plot is to be compared with the analogous plot in Fernández-Sánchez and Casanellas (2016); we summarize the comparison to other methods in Table 4.3. As it is apparent from the results in the table, SAQ outperforms all the other methods (even Erik+2 (m=2)) despite mixture data violates the assumptions of this method.

Performance of different methods applied to mixture data

internal branch length	SAQ	Erik+2 (m=2)	MP	ML(GTR+2 Γ)
0.01	**37**	12	0	0
0.05	**83**	35	2	4
0.1	**96**	60	19	14
0.2	**100**	86	76	77
0.3	**100**	96	99	95

TABLE 4.3: Percentage of correctly reconstructed topologies by different methods on alignments of length 1 000 bp for data generated under the GM model with 2 categories according to Section 4.1.1. The methods applied to reconstruct the topology were SAQ, Erik+2 with 2 partitions, Maximum Parsimony and ML(GTR+2 Γ) estimating time-reversible model with 2 discrete-gamma categories. In each row of the table, the highest success is indicated in bold font. In all the cases, ML had to estimate all parameters.

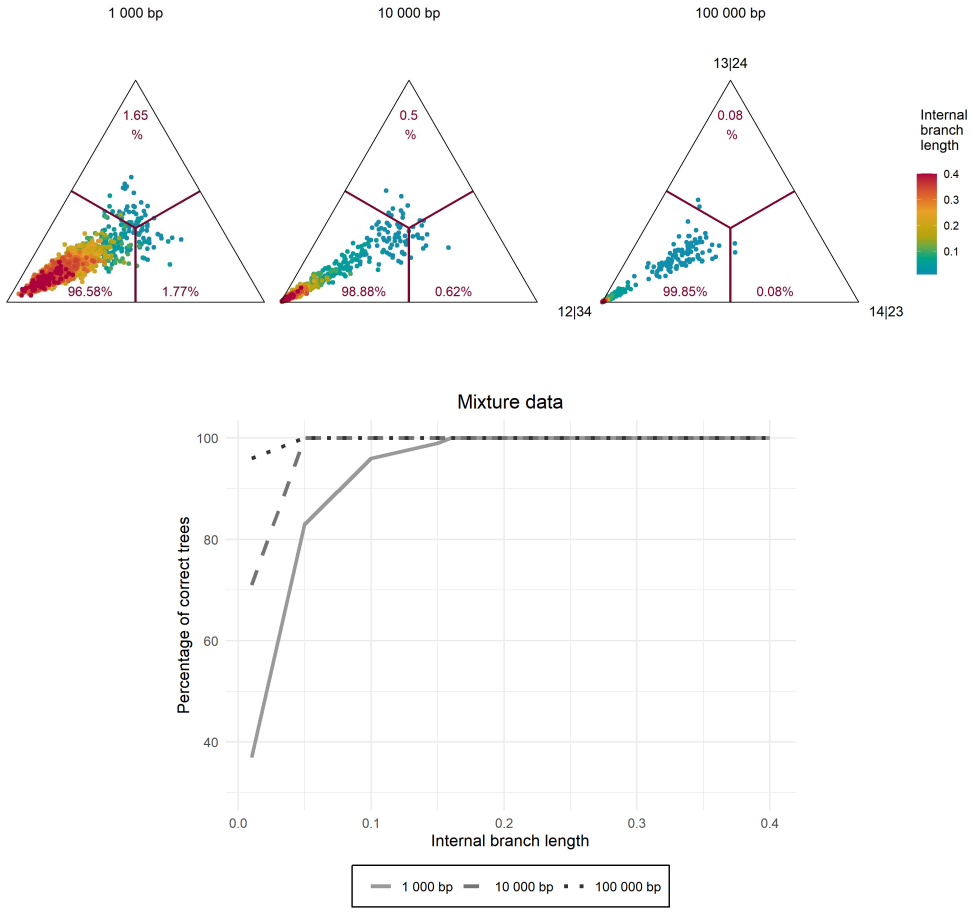


FIGURE 4.9: The three plots on the top show the ternary diagrams corresponding to the weights of SAQ applied to mixture data of lengths 1 000, 10 000 and 100 000 bp generated by the trees of Figure 4.2. On each triangle the bottom-left vertex represents the underlying true topology 12|34, the bottom-right vertex is topology 13|24 and the top vertex is 14|23. Each triangle is divided into three regions according to which tree is selected by the method. The figures in red represent the percentage of alignments of the corresponding region according to SAQ. The different colors of the dots represent the branch lengths r of the interior edge in the tree that gave rise to that dot. For the same data, the plot on the bottom represents the percentage of correctly reconstructed trees by SAQ as a function of the internal branch length for different length alignment. Data generated under the GM model with 2 categories, varying the internal branch length, and recovering with SAQ.

4.2.4 SAQ on real data

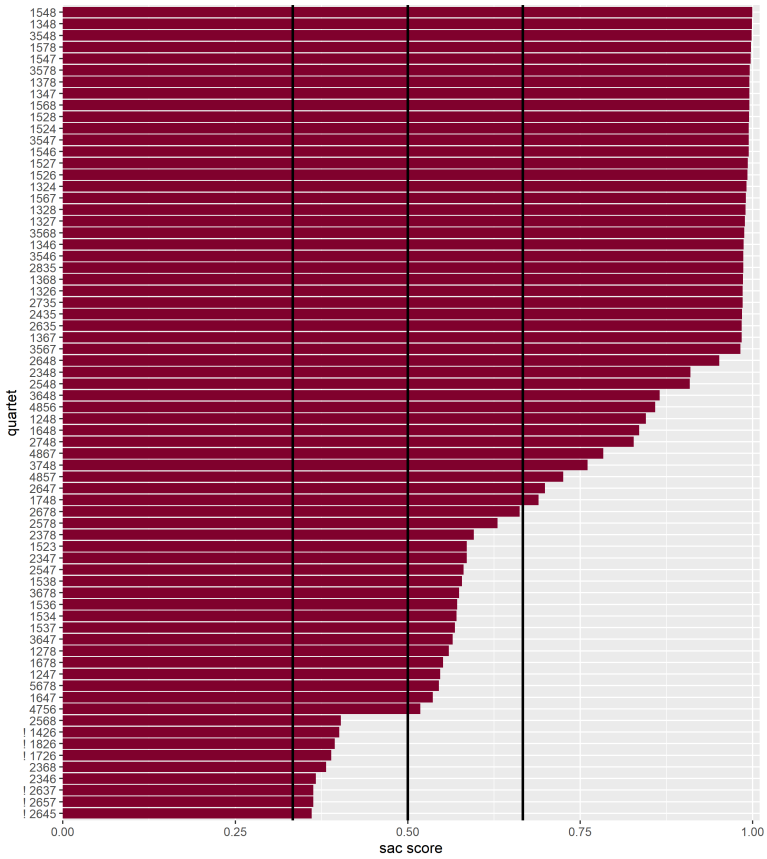


FIGURE 4.10: Performance of SAQ on real data set obtained from the yeast genome described in Section 4.1.3. Each number on the splits presented at the y axis correspond to a specie of yeast: 1) *C. albicans*, 2) *S. bayanus*, 3) *S. castellii*, 4) *S. cerevisiae*, 5) *S. kluyveri*, 6) *S. kudriavzevii*, 7) *S.mikatae*, 8) *S. paradoxus*. For each subalignment of 4 sequences it is presented the weight of SAQ for the quartet with maximum weight, are represented by the length of the horizontal dark bars. The 70 subalignment of 4 sequences have been ordered according to the highest weight of SAQ. The 6 quartets with the "!" symbol represent where the method has chosen quartet topology inconsistent with T in Figure 4.5. The vertical lines are placed at 0.33, 0.5 and 0.66.

We have also tested SAQ with the real data set presented in Section 4.1.3. In Figure 4.10 we present the highest score provided by SAQ on each of the 70 subalignments of 4 sequences. SAQ recovers correctly 91.42% of the underlying quartets on T of Figure 4.5. However, the quartets wrongly reconstructed (indicated with the the "!" symbol in the figure) represent compatible splits to T' in Figure 4.5.

4.3 PERFORMANCE OF THE PARALINEAR METHOD

In this section we include the results of assessing the performance of the paralinear method. Given a vector F of relative frequencies from an alignment, the method outputs the tree $A|B$ with highest score $I_{A|B}(F)$ (see Section 3.4). We have tested this method on the tree space with GM data as detailed in Section 4.1.1. The results can be found in Figure 4.11.

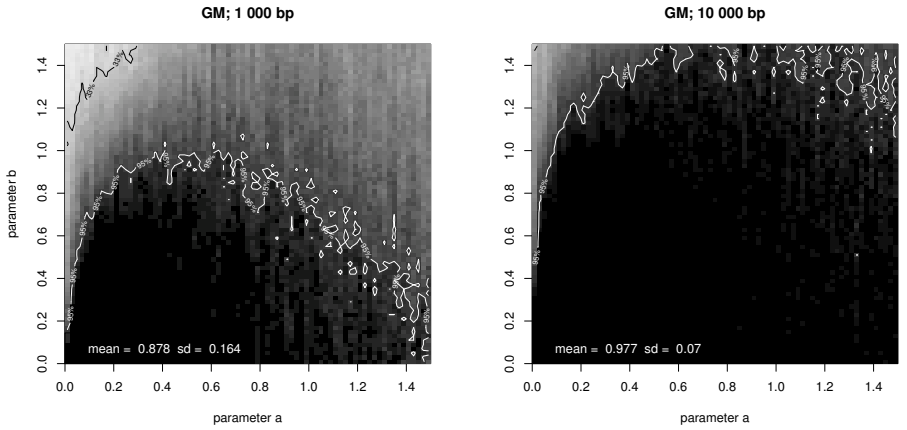


FIGURE 4.11: Performance of the paralinear method in the tree space of Figure 4.1 (b) on alignments of length 1 000 bp (left), 10 000 bp (right). Black is used to represent 100% of successful quartet reconstruction, white to represent 0%, and different tones of gray the intermediate frequencies. The 95% contour line is drawn in white, whereas the 33% contour line is drawn in black. Data were generated according to the GM model.

In Figure 4.12 we show the average value of $I_{A|B}(F)$ for a subset of alignments corresponding to trees in the tree space. Each plot corresponds to trees with fixed value of a and with b varying from 0.01 to 1.5. For each a, b one hundred alignments were considered. Note that for small values of a, $\frac{1}{8}I_{A|B}$ approximates the branch length of the interior edge of the tree, which follows from (1.12) and (3.12) (assuming that the distribution at the root is close to the uniform distribution). This can be observed in Figure 4.12. We can also observe that $I_{13|24}$ and $I_{14|23}$, which corresponds to the incompatible splits, have a similar behaviour (as expected, see Theorem 3.4.5 (b)) except when dealing with trees on the Felsenstein zone (that is small values of a and large values of b). In these cases $I_{13|24}$ is usually positive which leads to a bad performance of the method (see Figure 4.11).

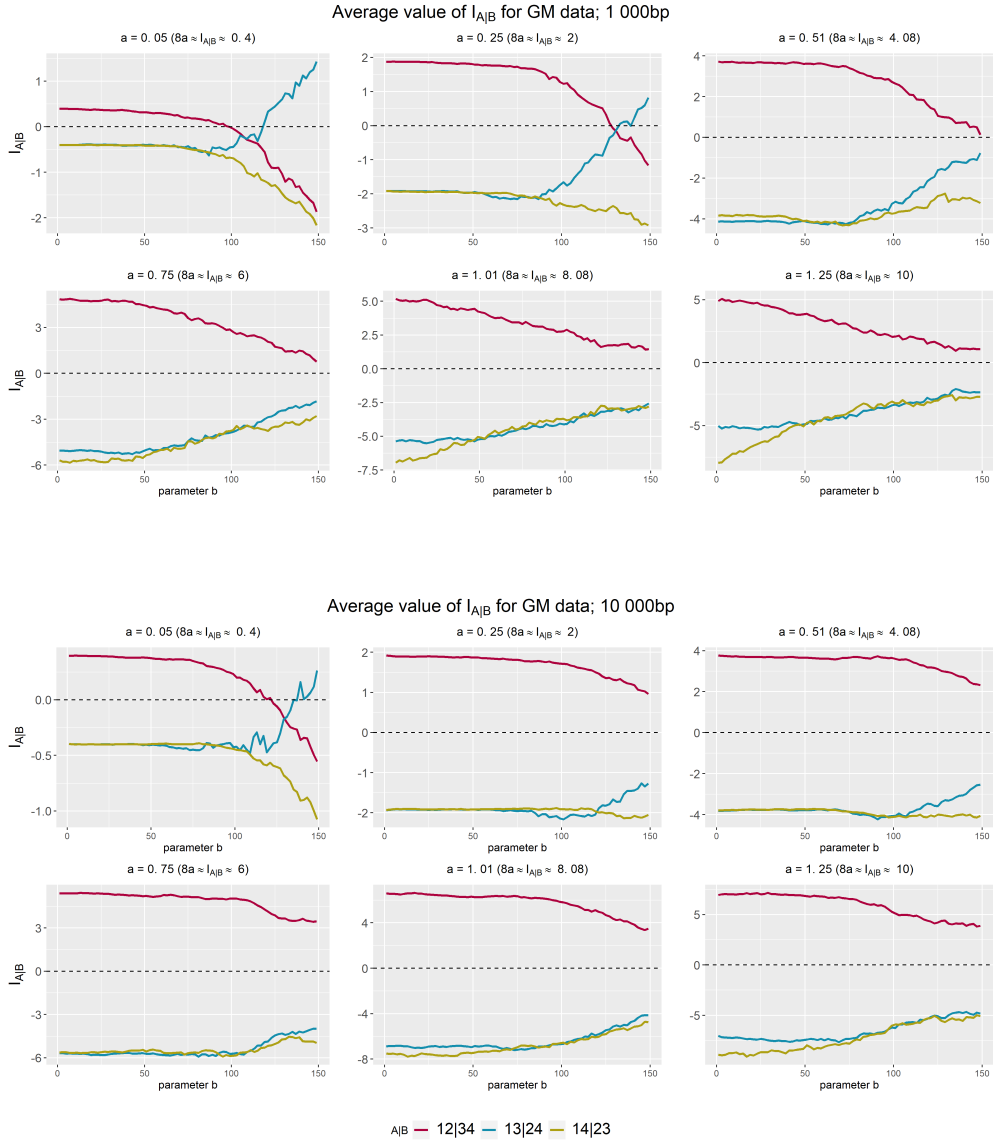


FIGURE 4.12: Average value of $I_{A|B}(F)$ for the 100 alignments of trees in the tree space of Figure 4.1 (b). Each plot correspond to a fixed value of the branch length a. The x axis corresponds to the value of b and the y axis to the avegare of $I_{A|B}(F)$. The top eight plots correspond to alignments of length 1 000 bp and the eight bottom plots to alignments of length 10 000 bp.

We see that the paralinear method is highly successful on the tree space with GM data and even outperforms SAQ. However, the method ASAQ, which makes use of the paralinear method in order to combine the methods Erik+2 and SAQ, achieves better results as we shall see in Section 4.4.1. On top of that, with ASAQ one can use the normalized weighting system of Erik+2 or SAQ for quartets in order to use a Q-method and reconstruct the topology of larger trees.

4.4 Performance of ASAQ on quartets

In this section we present the performance of ASAQ on the different scenarios of quartet data presented in the Section 4.1.1.

4.4.1 ASAQ on the tree space

The performance of ASAQ in recovering the correct quartet on data generated on the tree space of Section 4.1.1 is represented in Figure 4.13. As we have noted with SAQ, we observe the usual decreasing performance at the Felsenstein zone (both for GM and GTR data) and a higher performance at the other zones. Here we can also see that the success of the method increases with the sample size as expected.

For completeness, the average performance of ASAQ on this tree space for different alignment lengths and underlying models is compared to other methods in Table 4.4: we include the average results of SAQ and Erik+2 (on which ASAQ is based) and NJ and ML as published in Fernández-Sánchez and Casanellas (2016).

Average success of different methods on the tree space.

simulations	base pairs	ASAQ	SAQ	Erik+2	NJ	ML
GM	500	**85.3**	84.6	72.4	72.5	72.1
	1 000	**90**	88.8	80.3	79.7	73.6
	10 000	**98.4**	96.8	97.1	94.3	75.4
GTR	500	79.9	78.4	74.8	72.9	**88.0**
	1 000	86.9	83.5	84.3	80.5	**93.4**
	10 000	97.9	94.5	**99.2**	94.5	98

TABLE 4.4: Average success of ASAQ corresponding to the tree space of Figure 4.1 (b), compared to the results of the performance of SAQ, Erik+2, NJ with paralinear distance and ML taken from Casanellas, Fernández-Sánchez, and Garrote-López (2020) and Fernández-Sánchez and Casanellas (2016). ML(homGMc) is applied when data are generated under a GM model, while ML estimates a homogeneous GTR model when data are generated under GTR, see Table 1 in Fernández-Sánchez and Casanellas (2016). In each row of the table, the highest success is indicated in bold font.

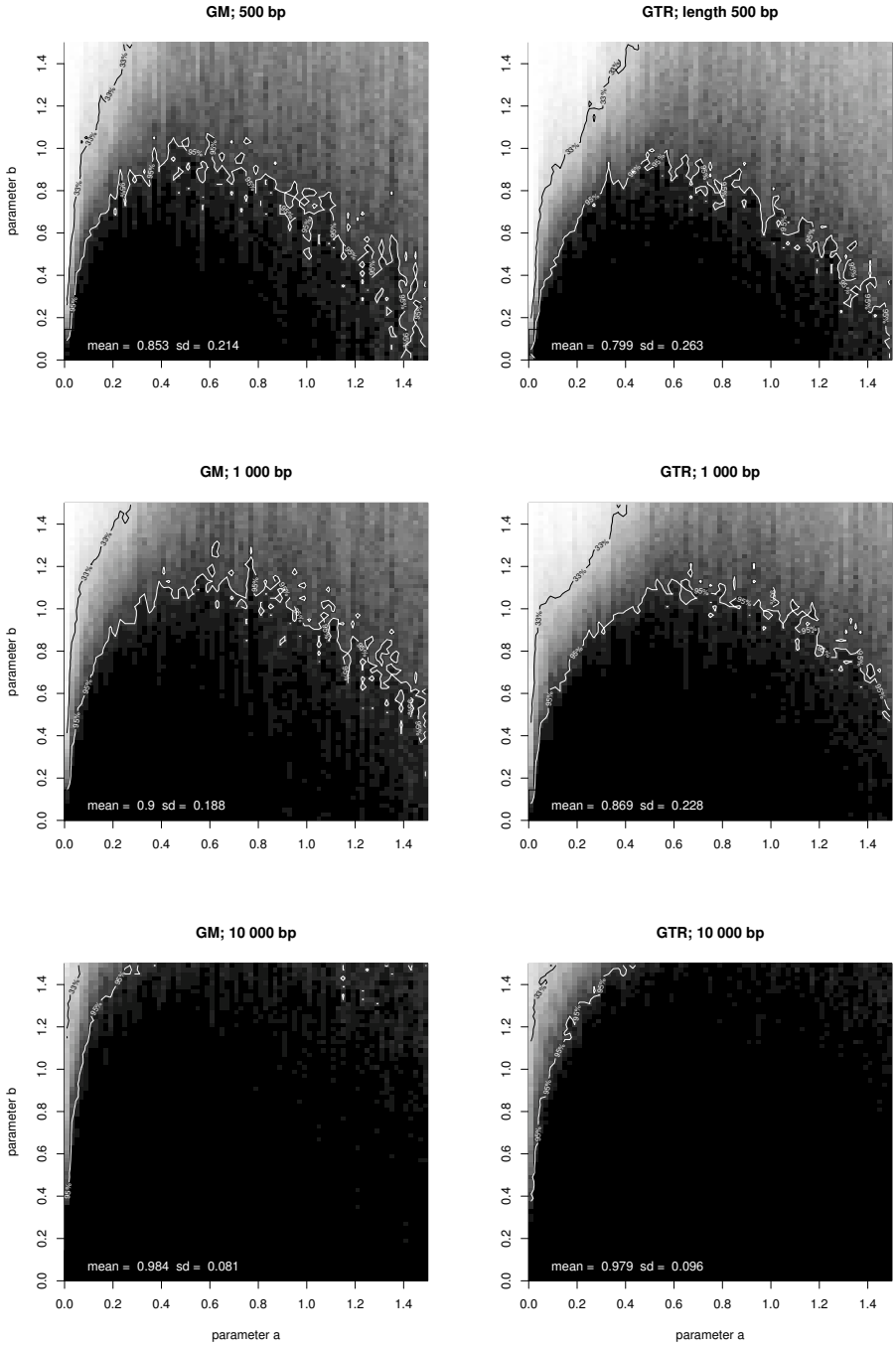


FIGURE 4.13: Performance of ASAQ in the tree space of Figure 4.1 (b) on alignments of length 500 bp (top), 1 000 bp (middle) and 10 000 bp (bottom). Black is used to represent 100% of successful quartet reconstruction, white to represent 0%, and different tones of gray the intermediate frequencies. The 95% contour line is drawn in white, whereas the 33% contour line is drawn in black. Left: data generated under the GM model; Right: data generated under a homogenous GTR model.

Table 4.4 shows that the performance of ASAQ is better than that Erik+2 (and SAQ) under GM data, specially for short alignments. For GTR data, the best perfomance for alignments of length 500 and 1 000 bp was that of ML (homGMc), while for long alignments (10 000 bp) the best results were obtained by Erik+2.

4.4.2 ASAQ on trees with random branch lengths

To visualize the overall distribution of the weights of ASAQ applied to trees with random branch lengths, we show ternary plots corresponding to the alignments previously described under the GM and GTR models and with branch lengths uniformly distributed in (0,1) (see Figure 4.14) and in (0,3) (see Figure 4.15).

Note that these ternary plots show a high performance of the method (see Table 4.5 for the summary of average success of ASAQ for these data), with a majority of points close to the left corner (which represents the correct quartet) and weights similarly distributed for the two discordant trees. This also shows that the method is not biased towards any of the incorrect topologies. We note that the level of success exhibited is quite sensitive to the branch length, being much higher for branch lengths in (0,1) than in (0,3). We do not appreciate a remarkable difference between the performance of ASAQ when applied to GTR data.

Average success of ASAQ applied to data generated on 12|34 with random branch lengths

branch length	GM		GTR	
	1 000 bp	10 000 bp	1 000 bp	10 000 bp
(0,1)	95.68	98.77	94.65	98.42
(0,3)	71.03	84.22	69.37	85.12

TABLE 4.5: Average success of ASAQ on alignments of lengths 1 000 and 10 000 bp generated on the tree 12|34 under the GM and GTR models with random branch lenghts uniformely distributed in (0,1) (first row) and (0,3) (second row). The plots corresponding to these data are shown in Figures 4.14 and 4.15.

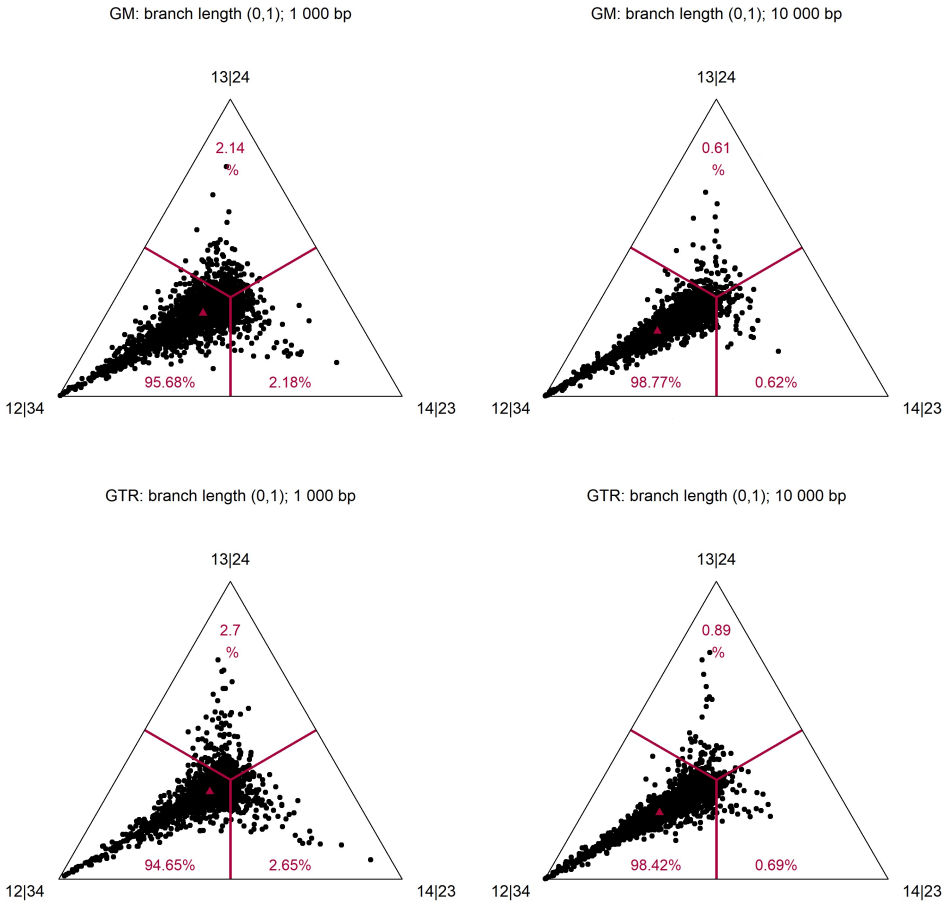


FIGURE 4.14: Ternary plots corresponding to the weights of `ASAQ` applied to 10 000 alignments generated under the 12|34 tree with random branch lengths uniformly distributed between 0 and 1. On each triangle the bottom-left vertex represents the underlying quartet 12|34, the bottom-right vertex is the quartet 13|24 and the top vertex is 14|23. The red triangle is the average of `ASAQ` weights applied to the 10 000 alignments. Top: correspond to data generated under GM; bottom: data generated under GTR. Left : 1 000 bp; Right: 10 000 bp.

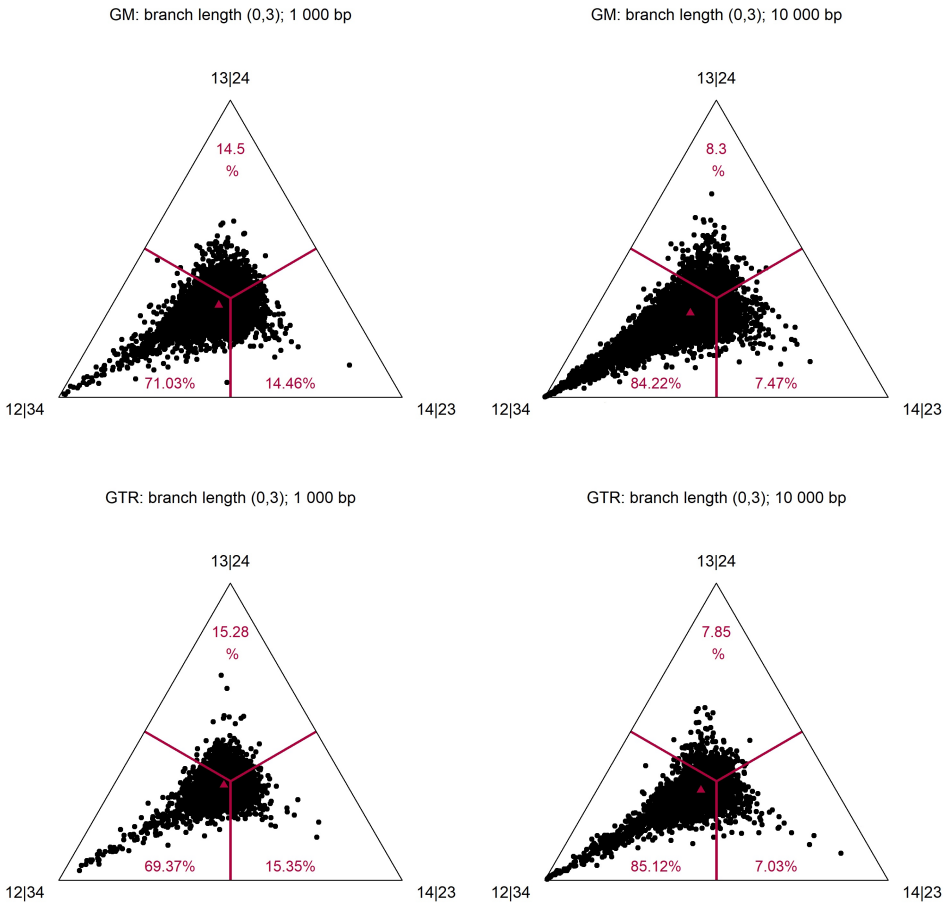


FIGURE 4.15: Ternary plots corresponding to the weights of ASAQ applied to 10 000 alignments generated under the 12|34 tree with random branch lengths uniformly distributed between 0 and 3. On each triangle the bottom-left vertex represents the underlying quartet 12|34, the bottom-right vertex is the quartet 13|24 and the top vertex is 14|23. The red triangle is the average of ASAQ weights applied to the 10 000 alignments. Top: correspond to data generated under GM; bottom: data generated under GTR. Left : 1 000 bp; Right: 10 000 bp.

4.4.3 ASAQ on mixture data

In Figure 4.16 we show the performance of ASAQ in determining the correct quartet when applied to mixture data. We also tested ASAQ with $m = 2$ meaning that it applies Erik+2 with $m = 2$ categories. Although we cannot prove the statistical consistency of ASAQ with $m = 2$, we analyzed the performance of this approach. Based on the comparison of Figure 5 of Fernández-Sánchez and Casanellas (2016), Figure 4.16 also compares this performance with that of Erik+2, Erik+2 ($m = 2$), maximum parsimony (MP) and two

versions of maximum likelihood. We did not include the performance of SAQ, as it is very similar to that of ASAQ (see Figure 4.9). The bar plots show an increasing accuracy of all the methods depending on the value of the parameter r. This was expected, as r accounts for the branch length at the interior edge of the two trees involved (see Figure 4.2) and larger r give larger divergence between sequences at the left and the right of the trees. We note that both ASAQ and ASAQ (m=2) (the first two bars of each plot) outperform the other methods, with a high level of success in all cases, even when the length of the alignments is 1 000 bp. The average success of ASAQ ($m = 2$) is 96.64% for 1 000 bp, 99% for 10 000 bp and 99.92% for 100 000 bp.

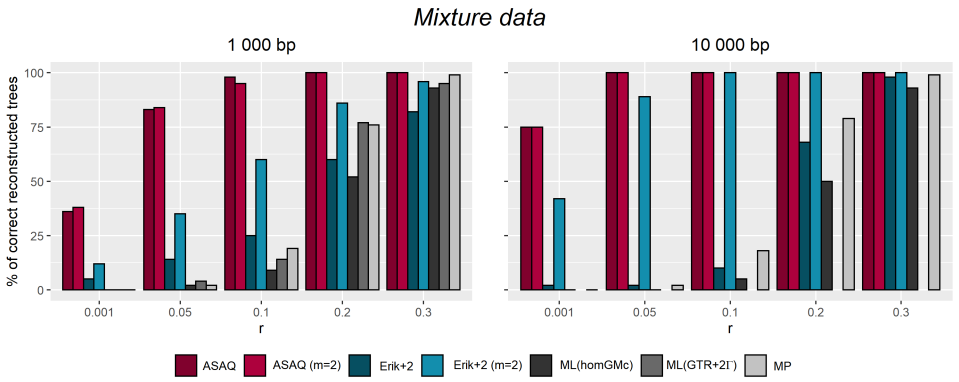


FIGURE 4.16: These bar plots represent the percentage of correctly reconstructed trees by several methods applied to the mixture data described in Section 4.1.1; the value r refers to the branch length of the interior edge of the trees (see Figure 4.2). The methods being compared are ASAQ, ASAQ ($m = 2$) and the methods presented in Figure 5 of Fernández-Sánchez and Casanellas (2016) on the same data: Erik+2, Erik+2 ($m = 2$), ML(homGMc) and a ML estimating a heterogeneous across lineages GTR with two categories of discrete Γ rates across sites, ML(GTR+2Γ) -only for length 1 000 bp- and maximum parsimony MP.

4.5 PERFORMANCE OF Q-METHODS

Since the output of quartet-based methods strongly depends on the choice of the initial quartet, each Q-method has been applied 100 times to each alignment and then the majority rule consensus tree (see Section 1.4.5) of these 100 replicates has been computed. We use the implementation of Q-methods on c++ provided by Marc Sabaté-Vidales (2014), and used on the quoted work. In order to evaluate the difference between two trees we use the Robinson-Foulds distance (RF for short, see Definition 1.1.13). For the computation of the majority rule consensus tree and the Robinson-Foulds distance with respect to an original tree, we have used the available functions in the Python Library *DendroPy*, see Sukumaran and Holder (2010).

4.5.1 Q-methods on general Markov data

In this section we present the results obtained by the Q-methods QP, WO and WIL applied to the input weights from ASAQ, Erik+2, NJ and ML on simulated GM data. These results are summarized in Figures 4.17 (for the tree topology *CC*), 4.18 (for *CD*) and 4.19 (for *DD*). The height of the bars show the average of the Robinson-Foulds distance from the original tree to the consensus tree of 100 replicates for each of the 100 generated alignments. We also plot in the figures the result of a global NJ applied with paralinear distance to the 12-sequence alignment; to this end we used the NJ algorithm implemented in the R package APE (Paradis, Claude, and Strimmer, 2004). These results are detailed in Tables B.1, B.2 and B.3 in Appendix B.

These results show a huge difference between the general performance of the quartet-based methods with input weights generated by ASAQ and Erik+2 methods against to the weights obtained by NJ and ML[1], specially when reconstructing the *CC* and *CD* trees. For these data, the Q-methods studied combined with Erik+2 or ASAQ weights clearly outperform the same methods with NJ or ML weights. This is specially true in the case of WO and WIL, which obtain the best results in general when combined with weights from ASAQ. The accuracy is not that good when applied to the *DD* tree: if the data were generated under GM, Erik+2 and ASAQ weights remain the best option combined with WO or WIL (for QP however, NJ weights are the best choice). All in all, the best reconstruction procedure seems to be the combination of WO with ASAQ weights. Applying WIL with ASAQ weights gives also very similar results.

For short alignments (600 bp) global NJ usually outperform all the other methods, except for $b \leq 0.015$ that has similar results to WO with ASAQ weights. However, for long alignments (5 000 and 10 000 bp) performance of global NJ and WO with ASAQ is very similar, indeed, WO+ASAQ gets slightly better results.

The reconstruction algorithms have had much more success in reconstructing the tree topology *CC*, than the *CD* and the *DD* tree topologies. Particularly, all reconstruction methods have a very bad performance when reconstructing the *DD* topology with GM data. In the tree topologies *CD* and *DD*, the distance between $seq9$ and $seq10$ is $4b$, while the distance between $seq8$ and $seq11$ is $20b$ (see Figure 4.4). The same happens between species $seq1$ and $seq2$, and $seq0$ and $seq3$ of the *DD* topology. Thus, when an alignment

[1]Note that ML did not converge for some quartets (specially in the presence of long branch attraction); in these cases we neglected the corresponding alignment and in the Appendix B tables, we denote between parentheses the number of alignments considered by ML. In those cases where ML did not converge for any of the 100 alignments, we have obviated the corresponding bar in the figures.

is generated on these trees, *long branch attraction* situation is created, which leads the reconstruction methods to infer a wrong topology.

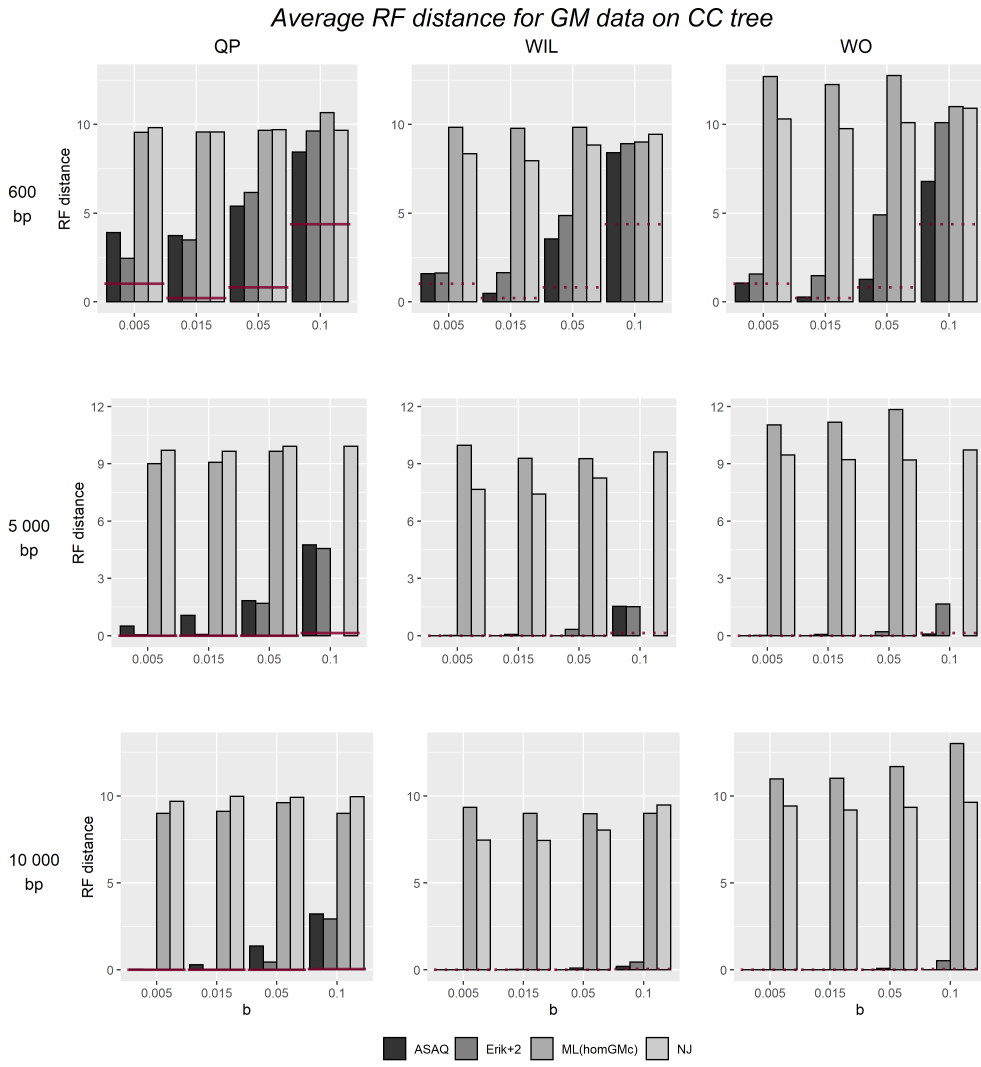


FIGURE 4.17: Average Robinson-Foulds distance for GM data simulated on the tree *CC* with alignment length 600 bp (above), 5 000 bp (middle) and 10 000 bp (below). The thick red line on the left (and dotted lines in the center and on the right) represent the average RF distance of the tree reconstructed using a global `NJ` with paralinear distance.

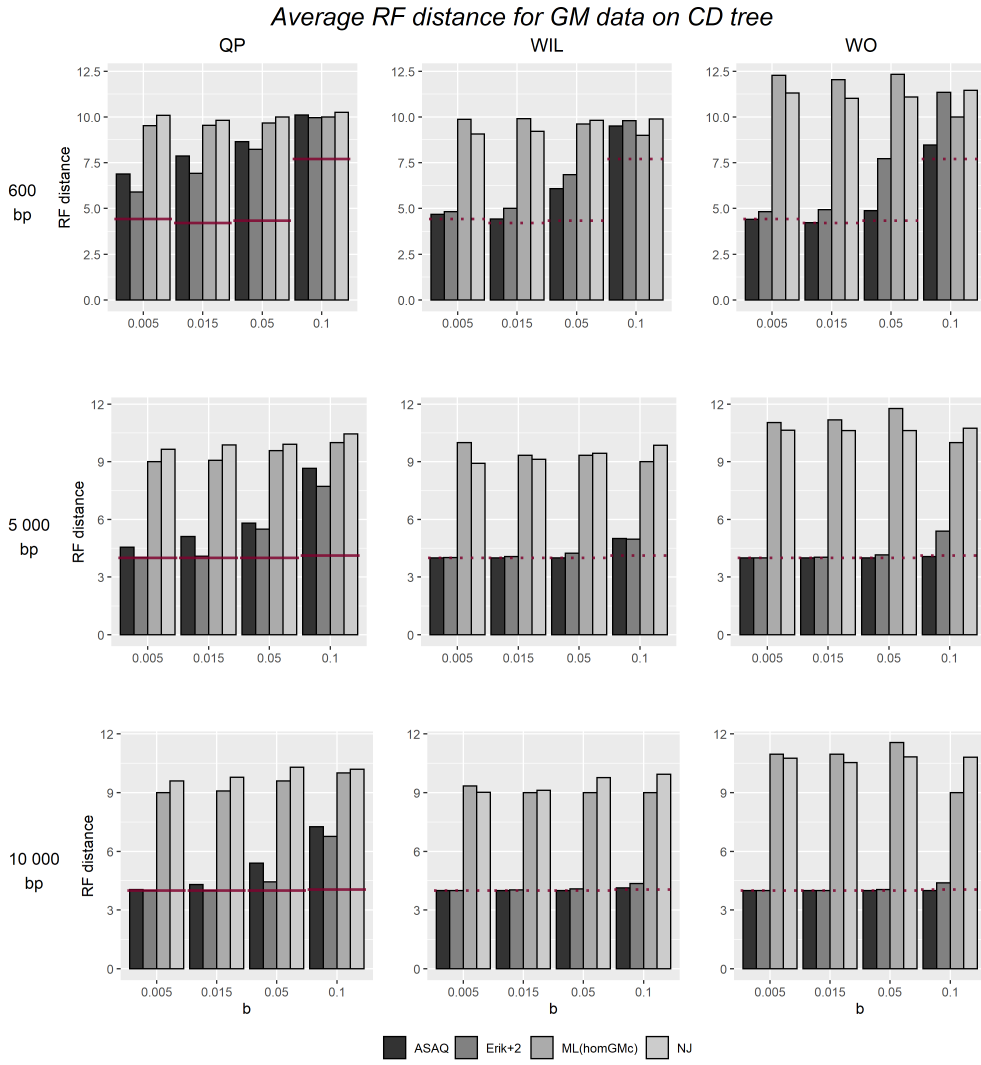


FIGURE 4.18: Average Robinson-Foulds distance for GM data simulated on the tree *CD* with alignment length 600 bp (above), 5 000 bp (middle) and 10 000 bp (below). The thick red line on the left (and dotted lines in the center and on the right) represent the average RF distance of the tree reconstructed using a global `NJ` with paralinear distance.

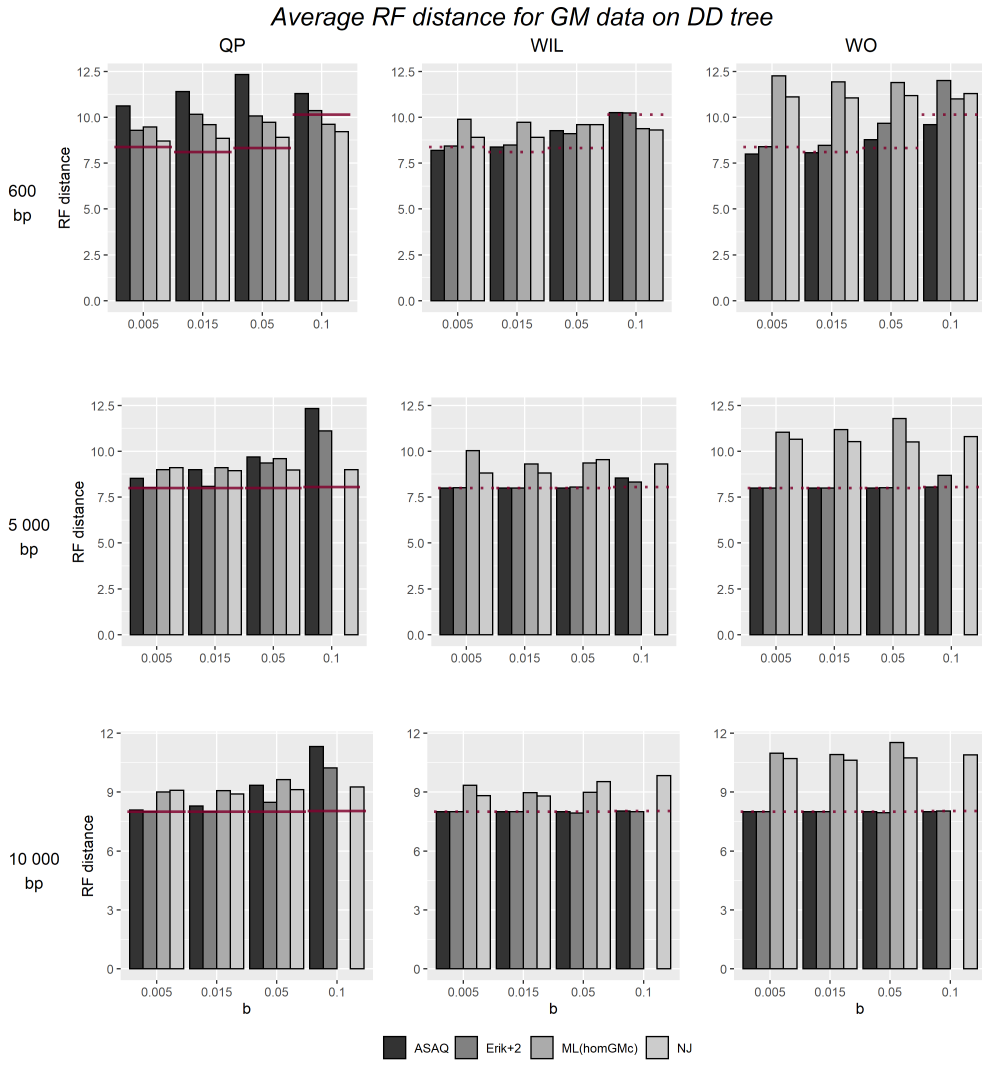


FIGURE 4.19: Average Robinson-Foulds distance for GM data simulated on the tree *DD* with alignment length 600 bp (above), 5 000 bp (middle) and 10 000 bp (below). The thick red line on the left (and dotted lines in the center and on the right) represent the average RF distance of the tree reconstructed using a global `NJ` with paralinear distance.

4.5.2 Q-methods on general time reversible data

When the data is generated under the GTR model, the combination of WO/WIL with `Erik+2/ASAQ` are still the best options, while the performance drops drastically when applied to short alignments (600 bp) generated from trees with long branch lengths (b=0.1).

The results obtained by QP, WO and WIL methods applied to the input weights of the different methods on GTR data are presented in Figure 4.20 and summarized in table B.4. This table corresponds to the performance of the methods considered over alignments of lengths 600 bp and 5 000 bp generated from the tree topology *DD* assuming the GTR model. We observe again a big difference in the performance of the quartet-based methods using `ASAQ` and `Erik+2` weights against that of `NJ` and `ML` weights. That is, the combination of WO/WILL with `Erik+2`/`ASAQ` are still the best options, while the performance drops drastically when applied to short alignments (600 bp) generated from trees with long branch lengths (b=0.1). It is remarkable that in these simulations, it is enough to consider alignments of length 5 000 bp to obtain an almost perfect reconstruction of the original tree using the WO method together with the `ASAQ` weights. It is also remarkable that for this particular topology the results described in the previous section, where the general Markov model was assumed, were not this good, obtaining distance values around 8 for the *DD* topology with 5 000bp.

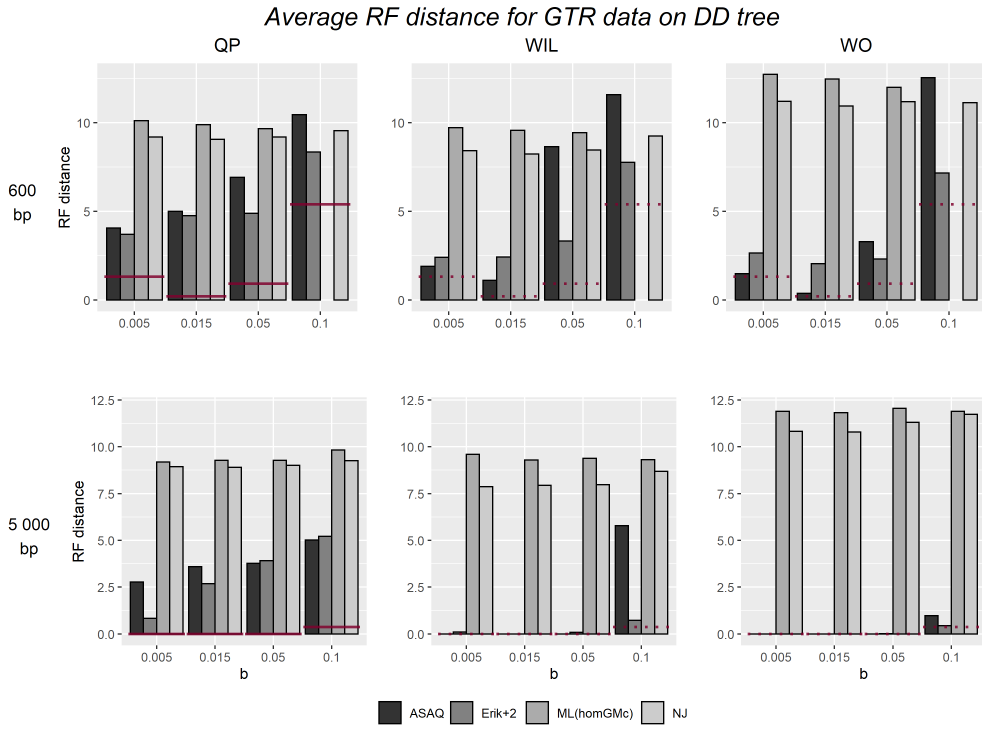


FIGURE 4.20: Average Robinson-Foulds distance for GTR data simulated on the tree *DD* tree. Alignment length 600 bp (above) and 5 000bp (below). The thick red line on the left (and dotted lines in the center and on the right) represent the average RF distance of the tree reconstructed using a global `NJ` with paralinear distance.

On GTR data, global NJ is the method that provide best results for alignments of length 600 bp. For larger alignments, the results of global NJ are very similar to WO+ASAQ (except for $b = 0.1$) and to WO+Erik+2.

4.5.3 Q-methods on mixture data

The results obtained by QP, WO and WIL applied to the input weights by ASAQ adapted or not to 2-category data ($m = 2$), Erik+2 with $m = 2$, NJ and ML on mixture data are summarized in Figure 4.21 (for QP method), 4.22 (for WIL method) and 4.23 (for WO method) and Tables B.5 to B.7. These figures and tables correspond to the performance of the methods over alignments of lengths 600, 5 000 and 10 000 generated from the tree topology *CD* assuming the 2-category model described in section 4.1.2. Tables B.5 to B.7 have the same structure as Tables B.1-B.4, and they correspond to the results obtained for different proportion between the two categories; $p = 0.25$, $p = 0.5$ and $p = 0.75$, respectively. As noted for the unmixed data, there is a big difference between the performance of the Q-methods applied with ASAQ and Erik+2 weights against their performance applied with NJ and ML.

It is remarkable that the reconstruction methods applied to mixture data on the *CD* tree is more accurate in average than when applied to unmixed data (Figure 4.18). This is probably due to the fact that when the branch lengths of species $seq3$, $seq4$ and $seq7$, $seq8$ are exchanged (see Figure 4.3), reconstruction methods have an easier job to make the appropriate splits, since the *long branch attraction* situation that was caused by the quartet of species $\{seq8, seq9, seq10, seq11\}$ does not longer exist with the new branch lengths. This is consistent with the observation that the reconstruction results are more accurate for low values of p (we recall that the proportion p of sites of the first category of the alignment were generated assuming the branch lengths of the *CD* tree in Fig. 4.3). As expected, the length of the alignment improves the performance of these methods, reducing the impact of the *long branch attraction* effect in the first category.

For these data, WO with ASAQ ($m = 2$) generally outperforms all other methods, including global NJ. For long alignment lengths (5 000 and 10 000 bp) the results of WO+ASAQ (either with $m = 2$ or not) and WO+Erik+2 ($m = 2$) also improve those of the global NJ. On the other hand, global NJ performs, in general, better than WIL and QP for any system of weights.

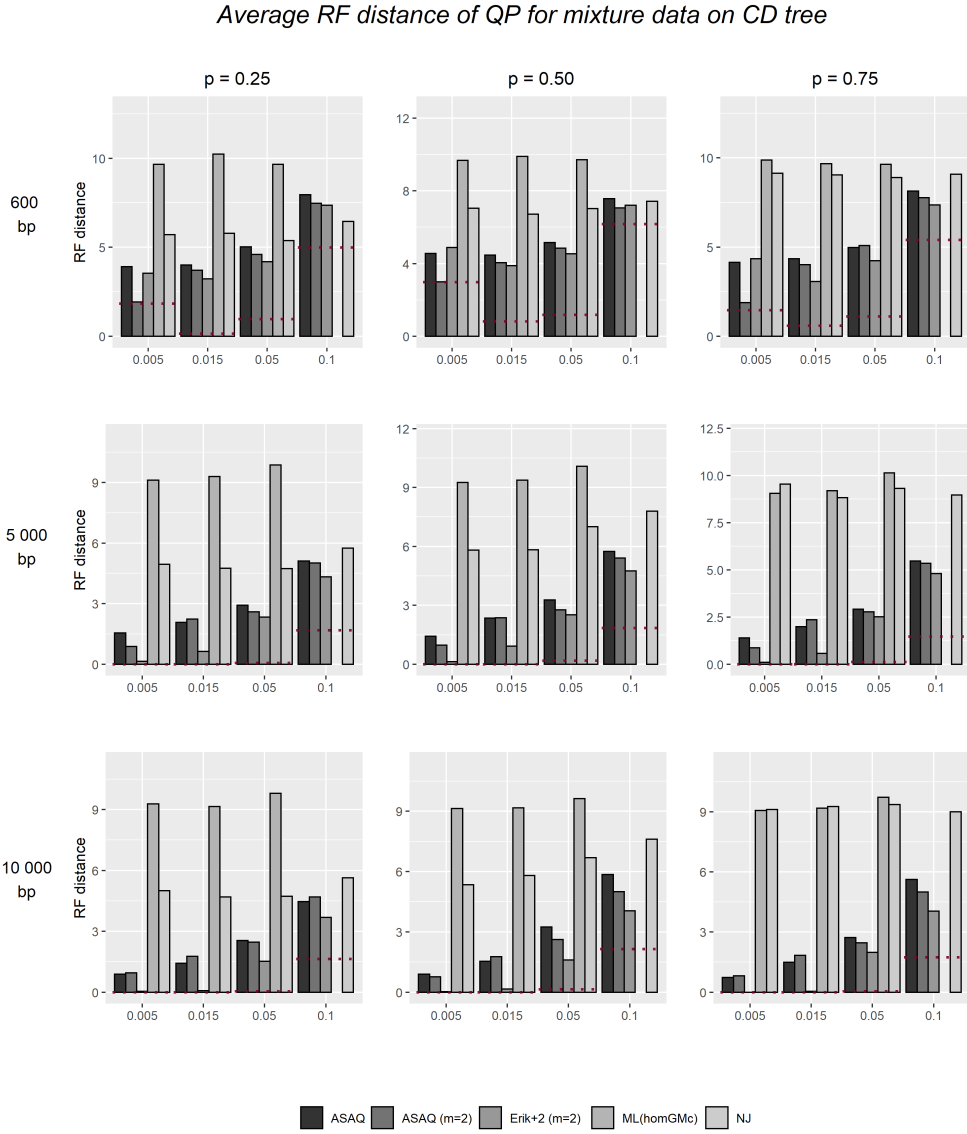


FIGURE 4.21: Results of the QP method on 2-category data with topology *CD* for $p = 0.25$ (top), $p = 0.5$ (middle) and $p = 0.75$ (bottom). The horizontal dotted lines are the result of a global `NJ` with the paralinear distance.

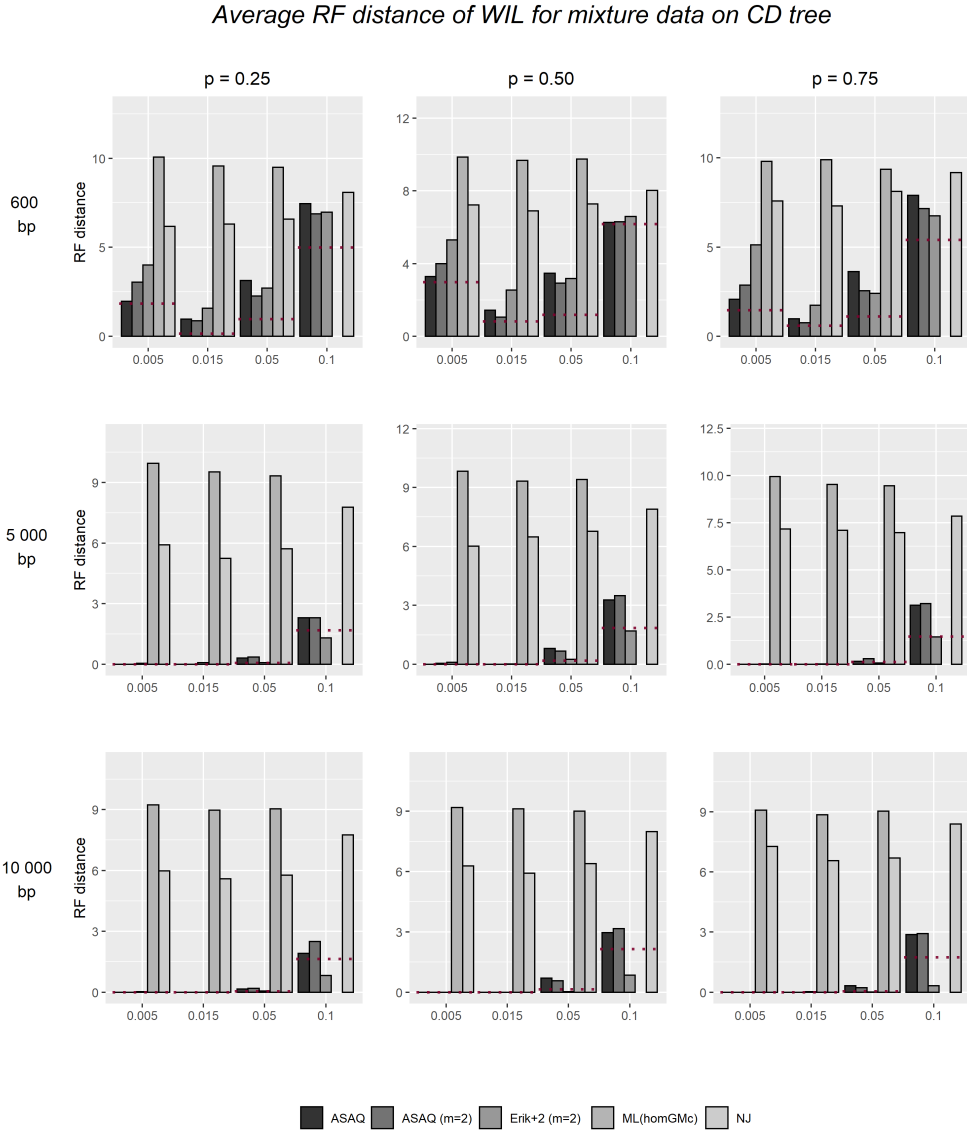


FIGURE 4.22: Results of the WIL method on 2-category data with topology *CD* for $p = 0.25$ (top), $p = 0.5$ (middle) and $p = 0.75$ (bottom). The horizontal dotted lines are the result of a global `NJ` with the paralinear distance.

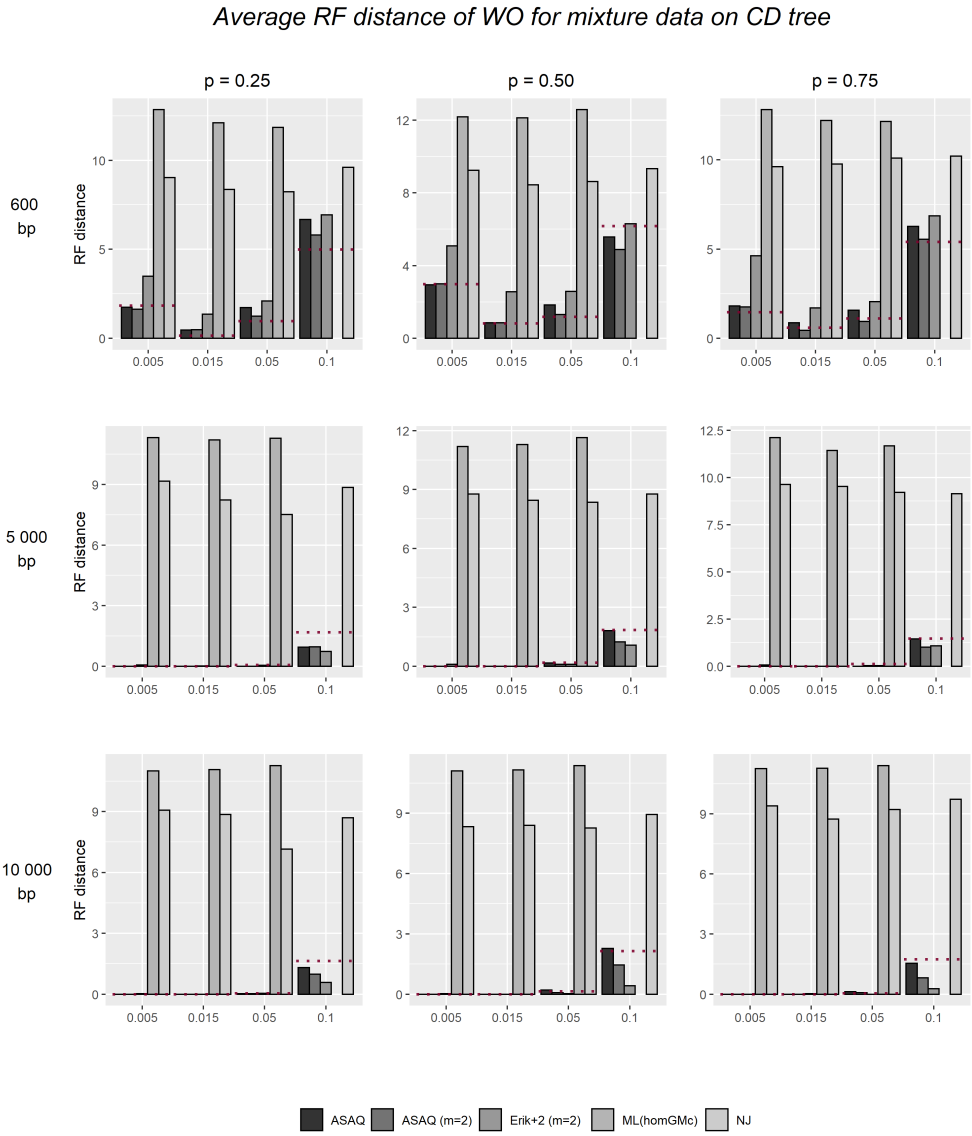


FIGURE 4.23: Results of the WO method on 2-category data with topology *CD* for $p = 0.25$ (top), $p = 0.5$ (middle) and $p = 0.75$ (bottom). The horizontal dotted lines are the result of a global `NJ` with the paralinear distance.

4.5.4 Q-methods on real data

We analyse the performance of ASAQ on real data on the eight species of yeast presented in Section 4.1.3. We investigate whether the quartets output by ASAQ support the tree T of Rokas et al. (2003) in Figure 4.5 left or the alternative tree T' of Phillips, Delsuc, and Penny (2004) (Figure 4.5 right).

According to Jayaswal et al. (2014) these data are best modeled by considering, apart from heterogeneity across lineages, two different rate categories (discrete Γ distribution) plus invariable sites. In our setting, this is translated into a mixture distribution with 3 categories ($m = 3$ in ASAQ). As one can see in Table 4.6, WO+ASAQ correctly reconstructs T when 3 categories are considered, but it reconstructs T' otherwise. Note that the RF distance of the consensus tree obtained by WIL+ASAQ to the tree T is much smaller than the distance to T' in concordance with Jayaswal et al. (2014). When applied to these data, a global NJ reconstructs the tree T.

	RF distance to T		RF distance to T'	
	WIL	WO	WIL	WO
ASAQ (m=1)	1.974	2	0.026	0
ASAQ (m=2)	1.922	2	0.078	0
ASAQ (m=3)	0.004	**0**	1.996	2

TABLE 4.6: Robinson-Foulds distance of the consensus tree obtained by WILL and WO with ASAQ weights (with different number of categories) to the trees T and T' suggested in Rokas et al. (2003) and Phillips, Delsuc, and Penny (2004), respectively.

4.6 DISCUSSION AND FURTHER CONSIDERATIONS

SAQ and ASAQ are two new methods of phylogenetic reconstruction that assume the most general nucleotide substitution model (a general Markov model) with independently and identically distributed sites. These methods are essentially based on the results presented in Theorem 1.3.24 (Allman and Rhodes, 2003 and 2008) and Theorem 1.3.26 (Allman, Rhodes, and Taylor, 2014) adapted to non-theoretical data as explained in Chapter 3, and take advantage of the stochastic information available in the data to infer the topology of the phylogenetic tree. The simulation studies performed show that SAQ and ASAQ are two robust and powerful methods for quartet topology reconstruction, especially when dealing with data generated under the GM model. Moreover, up to our knowledge, SAQ

is the first method that combines both the algebraic and semialgebraic nature of the underlying substitution models. We want to recall that since these methods are based on the algebraic and semialgebraic description of the model, they do not need to estimate the substitution parameters and in this sense, they may be adaptable to deal with amino acid substitution models.

Both, `SAQ` and `ASAQ` are developed for quartets and assign a normalized weight between 0 and 1 to each possible quartet: this weight should be the greatest for the quartet with highest confidence. These methods are statistically consistent, and the simulation studies seem to indicate that there is no tendency on these normalized weights towards any of the incorrect quartets. However, we have not computed the bias of the methods and we do not claim that they are asymptotically efficient (minimal variance); they may converge slowly under special situations. The simulations provided in this chapter are intended to overcome this lack of theoretical computations on the sample size required to ensure convergence with a certain probability. We are aware that we have not made a study of the method from the statistical point of view and we based the method on the asymptotics and the results on simulations. Such an study would certainly be relevant.

We have shown that the performance of the methods is good even for certain data sets that violate the model hypotheses. Indeed, the performance of both `SAQ` and `ASAQ` superior to other methods even if they are not proved to be statistically consistent on these data. Still, in Section 3.2 we have seen that `SAQ` can deal with some types of mixtures such as those presented in Proposition 3.2.15 and 3.2.18. Regarding the performance with mixture data, we would like to point out that, although these methods are not specifically designed to deal with mixtures of distributions in general, they can be easily adapted to deal with invariable sites. One only needs to redefine the relative frequency vector p taking into account the estimated number of invariable sites (e.g., as in Jayaswal, Robinson, and Jermiin, 2007 or Steel, Huson, and Lockhart, 2000). Although we do not present this implementation in this thesis, we plan to incorporate this feature in a future version of the software.

It is also worth noting that the mixture model that can be considered in `Erik+2` (and thus in `ASAQ`) is actually more flexible than we have mentioned. Indeed, the rank conditions considered in `Erik+2` for quartets remain valid if we let the substitution parameters of one cherry change in any number of categories (while the other cherry remains fixed in all the categories), see Casanellas and Fernández-Sánchez (2020). This makes the mixture model allowed by `Erik+2` and `ASAQ` more general, with implications for the performance of Q-methods with `ASAQ` weights on mixture data. However, we should study the statistical consistency of the methods in these cases with special care.

The results described in section 4.2 show that SAQ performs better when applied to data generated under the GM model than under a homogeneous GTR model (for which the results are satisfactory and similar to the Erik+2 method). One possible explanation for this phenomenon, is that when the transition matrix at the inner edge is close to the identity (so that the method is pushed to the limit), the flattening matrices $Flat_{12|34}$ (needed to compute the SAQ weights) are close to being symmetric and positive definite (at least when the stationary distribution is uniform) and this is more evident in GTR data because there is less variation in these data (e.g embeddable matrices only account for less than 0.06% of all transition matrices, see Roca-Lacostena, 2021). In this situation, the semialgebraic information of the data becomes almost irrelevant and the method is practically based on the distance to the 4-rank matrices. As a consequence, the results are similar to those of Erik+2 and other methods based on rank conditions. Still ASAQ manages to improve the results of SAQ and Erik+2 in this case, especially for short lengths in the alignments.

We have seen that WO with ASAQ weights is a phylogenetic reconstruction method with compatible performance to that of a global NJ, which overpasses NJ in some situations, as when applied to mixture data. This contrasts with the results obtained in Ranwez and Gascuel (2001) and St. John et al. (2003), where only weights from ML and NJ were considered. The results in this thesis validate the Q-method WO as a good phylogenetic reconstruction method if it is applied together with a system of reliable weights as suggested in Ranwez and Gascuel (2001). Moreover, as ASAQ assumes the most general Markov model and can deal with mixtures, this opens the door to use Q-methods under complex models. Note than WO has a higher order of computational complexity than NJ ($O(n^4)$ against $O(n^3)$), but we have not explored the possibility of considering a subset of the possible 4-tuples as suggested in Snir and Rao (2010) or Davidson et al. (2018). It would be interesting to consider these powerful methods with weights from ASAQ or to restrict to quartets with highest weights as starting point for Q-methods. We have not considered a comparison to a global maximum likelihood because it seems to have a similar performance to NJ in practice (Ranwez and Gascuel, 2001), and in St. John et al. (2003) it is claimed that NJ usually outperforms ML.

Finally, let us note that combining algebraic and semi-algebraic conditions in a single score has been a hard task. There are many different ways of combining all the restrictions satisfied by the theoretical distributions. For example, in the definition of the score s_T^i of the SAQ method, we could have considered not projecting into symmetric positive matrices for the flattening $Flat_{12|34}$ as this seems as adding unnecessary complexity to the algorithm. However, we found that by considering the projection, all distances to

4-rank matrices are computed on the same space of positive definite matrices and this homogeneity improves the results of the algorithm. We have tested many other options that are not mentioned in this final memoir. In this sense, `SAQ` and `ASAQ` are the result of a number of decisions based on exhaustive simulation studies, with the scope of obtaining a method that takes into account all the information at hand, but at the same time, keeping the method computationally feasible and as simple as possible.

Bibliography

Allman, Elizabeth S., James H. Degnan, and John A. Rhodes (2013). "Species Tree Inference by the STAR Method and Its Generalizations". In: *Journal of Computational Biology* 20(1), pp. 50–61. ISSN: 1066-5277. DOI: 10.1089/cmb.2012.0101.

Allman, Elizabeth S., Laura S. Kubatko, and John A. Rhodes (2016). "Split Scores: A Tool to Quantify Phylogenetic Signal in Genome-Scale Data". In: *Systematic Biology* 66(4), syw103. ISSN: 1063-5157. DOI: 10.1093/sysbio/syw103.

Allman, Elizabeth S. and John A. Rhodes (2003a). *Mathematical Models in Biology*. Cambridge University Press. ISBN: 9780521525862. DOI: 10.1017/CBO9780511790911.

Allman, Elizabeth S. and John A. Rhodes (2003b). "Phylogenetic invariants for the general Markov model of sequence mutation". In: *Mathematical Biosciences* 186(2), pp. 113–144. ISSN: 00255564. DOI: 10.1016/j.mbs.2003.08.004.

Allman, Elizabeth S. and John A. Rhodes (2004). "Quartets and Parameter Recovery for the General Markov Model of Sequence Mutation". In: *Applied Mathematics Research eXpress* 2004(4), pp. 107–131. ISSN: 1687-1200. DOI: 10.1155/S1687120004020283.

Allman, Elizabeth S. and John A. Rhodes (2005). *The mathematics of Phylogenetics*. University of Alaska Fairbanks.

Allman, Elizabeth S. and John A. Rhodes (2006). "The Identifiability of Tree Topology for Phylogenetic Models, Including Covarion and Mixture Models". In: *Journal of Computational Biology* 13(5), pp. 1101–1113. DOI: 10.1089/cmb.2006.13.1101.

Allman, Elizabeth S. and John A. Rhodes (2008). "Phylogenetic ideals and varieties for the general Markov model". In: *Advances in Applied Mathematics* 40(2), pp. 127–148. ISSN: 01968858. DOI: 10.1016/j.aam.2006.10.002.

Allman, Elizabeth S., John A. Rhodes, and Amelia Taylor (2014). "A Semialgebraic Description of the General Markov Model on Phylogenetic Trees". In: *SIAM Journal on Discrete Mathematics* 28(2), pp. 736–755. ISSN: 0895-4801. DOI: 10.1137/120901568. eprint: 1212.1200 (q-bio.PE).

Altschul, Stephen F. et al. (1990). “Basic local alignment search tool”. In: *Journal of Molecular Biology* 215(3), pp. 403–410. ISSN: 00222836. DOI: 10.1016/S0022-2836(05)80360-2.

Barry, Daniel and John A. Hartigan (1987). “Asynchronous Distance between Homologous DNA Sequences”. In: *Biometrics* 43(2), p. 261. ISSN: 0006341X. DOI: 10.2307/2531811.

Bosma, Wieb, John Cannon, and Catherine Playoust (1997). “The Magma Algebra System I: The User Language”. In: *Journal of Symbolic Computation* 24(3-4), pp. 235–265. ISSN: 07477171. DOI: 10.1006/jsco.1996.0125.

Bray, Nicolas L. and Lior Pachter (2003). “MAVID multiple alignment server”. In: *Nucleic acids research* 31(13), pp. 3525–3526. ISSN: 1362-4962. DOI: 10.1093/nar/gkg623.

Buneman, Peter (1971). “The recovery of trees from measures of dissimilarity”. In: *Mathematics in the Archaeological and Historical Sciences*. Ed. by Edinburgh University Press, pp. 387–395.

Camin, Joseph H. and Robert R. Sokal (1965). “A Method for Deducing Branching Sequences in Phylogeny”. In: *Evolution* 19(3), p. 311. ISSN: 00143820. DOI: 10.2307/2406441.

Casanellas, Marta and Jesús Fernández-Sánchez (2007). “Performance of a New Invariants Method on Homogeneous and Nonhomogeneous Quartet Trees”. In: *Molecular Biology and Evolution* 24(1), pp. 288–293. ISSN: 1537-1719. DOI: 10.1093/molbev/msl153.

Casanellas, Marta and Jesús Fernández-Sánchez (2008). “Geometry of the Kimura 3-parameter model”. In: *Advances in Applied Mathematics* 41(3), pp. 265–292. ISSN: 01968858. DOI: 10.1016/j.aam.2007.09.003.

Casanellas, Marta and Jesús Fernández-Sánchez (2011). “Relevant phylogenetic invariants of evolutionary models”. In: *Journal de Mathématiques Pures et Appliquées* 96(3), pp. 207–229. ISSN: 00217824. DOI: 10.1016/j.matpur.2010.11.002.

Casanellas, Marta and Jesús Fernández-Sánchez (2020). “Rank conditions on phylogenetic networks”. In: *Research Perspectives CRM Barcelona. Spring 2019*. Vol. 10. Trends in Mathematics. Springer-Birkhäuser, to appear.

Casanellas, Marta, Jesús Fernández-Sánchez, and Marina Garrote-López (2020). “SAQ: semi-algebraic quartet reconstruction method”. In: *ArXiv 3488312*, pp. 1–28. arXiv: 2011.13968.

Casanellas, Marta, Jesús Fernández-Sánchez, and Anna M. Kedzierska (2012). “The space of phylogenetic mixtures for equivariant models”. DOI: 10.1186/1748-7188-7-33.

Casanellas, Marta, Jesús Fernández-Sánchez, and Mateusz Michałek (2015). "Low degree equations for phylogenetic group-based models". In: *Collectanea Mathematica* 66(2), pp. 203–225. ISSN: 0010-0757. DOI: 10.1007/s13348-014-0120-0.

Casanellas, Marta, Jesús Fernández-Sánchez, and Mateusz Michałek (2017). "Local equations for equivariant evolutionary models". In: *Advances in Mathematics* 315, pp. 285–323. ISSN: 00018708. DOI: 10.1016/j.aim.2017.05.003.

Cavender, James A. and Joseph Felsenstein (1987). "Invariants of phylogenies in a simple case with discrete states". In: *Journal of Classification* 4(1), pp. 57–71. ISSN: 0176-4268. DOI: 10.1007/BF01890075.

Chang, Joseph T. (1996). "Full reconstruction of Markov models on evolutionary trees: Identifiability and consistency". In: *Mathematical Biosciences* 137(1), pp. 51–73. ISSN: 00255564. DOI: 10.1016/S0025-5564(96)00075-2.

Chifman, Julia and Laura S. Kubatko (2014). "Quartet Inference from SNP Data Under the Coalescent Model". In: *Bioinformatics* 30(23), pp. 3317–3324. ISSN: 1460-2059. DOI: 10.1093/bioinformatics/btu530.

Chifman, Julia and Laura S. Kubatko (2015). "Identifiability of the unrooted species tree topology under the coalescent model with time-reversible substitution processes, site-specific rate variation, and invariable sites". In: *Journal of Theoretical Biology* 374, pp. 35–47. ISSN: 00225193. DOI: 10.1016/j.jtbi.2015.03.006.

Cleaves, Henderson J. (2011). "Watson–Crick Pairing". In: *Encyclopedia of Astrobiology*. Springer Berlin Heidelberg: Berlin, Heidelberg, pp. 1775–1776. DOI: 10.1007/978-3-642-11274-4_1683.

Cox, David A., John Little, and Donal O'Shea (2007). *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*. Third. Undergraduate Texts in Mathematics. Springer Publishing Company, Incorporated: New York, pp. xiv+536. ISBN: 978-0-387-35651-8.

Davidson, Ruth et al. (2018). "Efficient Quartet Representations of Trees and Applications to Supertree and Summary Methods". In: *IEEE/ACM Transactions on Computational Biology and Bioinformatics* 15(3), pp. 1010–1015. ISSN: 1545-5963. DOI: 10.1109/TCBB.2016.2638911.

Draisma, Jan and Jochen Kuttler (2009). "On the ideals of equivariant tree models". In: *Mathematische Annalen* 344(3), pp. 619–644. ISSN: 0025-5831. DOI: 10.1007/s00208-008-0320-6.

Draisma, Jan et al. (2016). "The Euclidean Distance Degree of an Algebraic Variety". In: *Foundations of Computational Mathematics* 16(1), pp. 99–149. ISSN: 1615-3375. DOI: 10.1007/s10208-014-9240-x.

Eckart, Carl and Gale Young (1936). "The approximation of one matrix by another of lower rank". In: *Psychometrika* 1(3), pp. 211–218. ISSN: 0033-3123. DOI: 10.1007/BF02288367.

Evans, Steven N. and T. P. Speed (1993). "Invariants of Some Probability Models Used in Phylogenetic Inference". In: *The Annals of Statistics* 21(1), pp. 355–377. ISSN: 0090-5364. DOI: 10.1214/aos/1176349030.

Fan, Ky and A. J. Hoffman (1955). "Some Metric Inequalities in the Space of Matrices". In: *Proceedings of the American Mathematical Society* 6(1), p. 111. ISSN: 00029939. DOI: 10.2307/2032662.

Felsenstein, Joseph (1973). "Maximum-likelihood estimation of evolutionary trees from continuous characters." In: *American journal of human genetics* 25(5), pp. 471–92. ISSN: 0002-9297.

Felsenstein, Joseph (1978). "Cases in which Parsimony or Compatibility Methods Will be Positively Misleading". In: *Systematic Zoology* 27(4), p. 401. ISSN: 00397989. DOI: 10.2307/2412923.

Fernández-Sánchez, Jesús and Marta Casanellas (2016). "Invariant Versus Classical Quartet Inference When Evolution is Heterogeneous Across Sites and Lineages". In: *Systematic Biology* 65(2), pp. 280–291. ISSN: 1063-5157. DOI: 10.1093/sysbio/syv086.

Foster, Peter G. (2004). "Modeling Compositional Heterogeneity". In: *Systematic Biology* 53(3). Ed. by Ted Schultz, pp. 485–495. ISSN: 1063-5157. DOI: 10.1080/10635150490445779.

Frobenius, Ferdinand G. (1912). *Über Matrizen aus nicht negativen Elementen*. Preussische Akademie der Wissenschaften Berlin: Sitzungsberichte der Preußischen Akademie der Wissenschaften zu Berlin. Reichsdr. ISBN: 9783111271903.

Galtier, Nicolas and Manolo Gouy (1998). "Inferring pattern and process: maximum-likelihood implementation of a nonhomogeneous model of DNA sequence evolution for phylogenetic analysis". In: *Molecular Biology and Evolution* 15(7), pp. 871–879. ISSN: 0737-4038. DOI: 10.1093/oxfordjournals.molbev.a025991.

Garrote-López, Marina (2016). "Stochasticity conditions for the general Markov model". In: *Reports@SCM* 2(1), pp. 45–57. DOI: 10.2436/20.2002.02.10.

Gascuel, Olivier (1994). "A note on Sattath and Tversky's, Saitou and Nei's, and Studier and Keppler's algorithms for inferring phylogenies from evolutionary distances." In: *Molecular Biology and Evolution* 11, pp. 961–963. ISSN: 1537-1719. DOI: 10.1093/oxfordjournals.molbev.a040176.

Ghys, Étienne and Andrew Ranicki (2016). "Signatures in algebra, topology and dynamics". In: *Ensaios Matemáticos* 30, pp. 1–173.

Grayson, Daniel R. and Michael E. Stillman (2009). *Macaulay2, a software system for research in algebraic geometry*. URL: http://www.math.uiuc.edu/Macaulay2/.

Greuel, Gert-Martin W., Gerhard Pfister, and Hans Schönemann (2009). *SINGULAR: a computer algebra system for polynomial computations*. DOI: 10.1145/1504347.1504377.

Gross, Elizabeth, Sonja Petrović, and Jan Verschelde (2013). "Interfacing with PHCpack". In: *Journal of Software for Algebra and Geometry* 5, pp. 20–25. DOI: 10.2140/jsag.2013.5.20.

Gross, Elizabeth et al. (2016). "Numerical algebraic geometry for model selection and its application to the life sciences". In: *Journal of The Royal Society Interface* 13(123), p. 20160256. ISSN: 1742-5689. DOI: 10.1098/rsif.2016.0256.

Gusfield, Dan (1997). *Algorithms on Strings, Trees and Sequences: Computer Science and Computational Biology*. Cambridge University Press: USA. ISBN: 9780521585194. DOI: 10.1017/CBO9780511574931.

Haeckel, Ernst Heinrich Philipp August (1866). *Generelle morphologie der organismen. Allgemeine grundzüge der organischen formen-wissenschaft, mechanisch begründet durch die von Charles Darwin reformirte descendenztheorie*. G. Reimer: Berlin. DOI: 10.5962/bhl.title.3953.

Hendy, Michael D. (1989). "The Relationship Between Simple Evolutionary Tree Models and Observable Sequence Data". In: *Systematic Zoology* 38(4), p. 310. ISSN: 00397989. DOI: 10.2307/2992397.

Hendy, Michael D. and David Penny (1989). "A Framework for the Quantitative Study of Evolutionary Trees". In: *Systematic Zoology* 38(4), p. 297. ISSN: 00397989. DOI: 10.2307/2992396.

Hendy, Michael D., David Penny, and Mike A. Steel (1994). "A discrete Fourier analysis for evolutionary trees." In: *Proceedings of the National Academy of Sciences* 91(8), pp. 3339–3343. ISSN: 0027-8424. DOI: 10.1073/pnas.91.8.3339.

Higham, Nicholas J. (1988). "Computing a nearest symmetric positive semidefinite matrix". In: *Linear Algebra and its Applications* 103, pp. 103–118. ISSN: 00243795. DOI: 10.1016/0024-3795(88)90223-6.

Higham, Nicholas J. (1989). "Matrix Nearness Problems and Applications". In: *Applications of Matrix Theory* 22.

Ho, Simon Y.W. and Lars S. Jermiin (2004). "Tracing the Decay of the Historical Signal in Biological Sequence Data". In: *Systematic Biology* 53(4). Ed. by Peter Lockhart, pp. 623–637. ISSN: 1076-836X. DOI: 10.1080/10635150490503035.

Horn, Roger A. and Charles R. Johnson (1985). *Matrix Analysis*. Cambridge University Press. ISBN: 9780521386326. DOI: 10.1017/CBO9780511810817.

Huelsenbeck, John P. (1995). "Performance of Phylogenetic Methods in Simulation". In: *Systematic Biology* 44(1), pp. 17–48. ISSN: 1076-836X. DOI: 10.1093/sysbio/44.1.17.

Jacobi, Carl G. J. (1857). "Ueber eine elementare Transformation eines in bezug auf jedes von zwei Variablen-Systemen linearen und homogenen Ausdrucks". In: *Journal für die reine und angewandte Mathematik* 53(3), pp. 583–590.

Jarvis, Peter D., Barbara R. Holland, and Jeremy G. Sumner (2013). "Phylogenetic Invariants". In: *Brenner's Encyclopedia of Genetics*. Elsevier: Oxford, pp. 321–323. DOI: 10.1016/B978-0-12-374984-0.00815-9.

Jayaswal, Vivek, John Robinson, and Lars S. Jermiin (2007). "Estimation of Phylogeny and Invariant Sites under the General Markov Model of Nucleotide Sequence Evolution". In: *Systematic Biology* 56(2). Ed. by Mike Steel, pp. 155–162. ISSN: 1076-836X. DOI: 10.1080/10635150701247921.

Jayaswal, Vivek et al. (2014). "Mixture Models of Nucleotide Sequence Evolution that Account for Heterogeneity in the Substitution Process Across Sites and Across Lineages". In: *Systematic Biology* 63(5), pp. 726–742. DOI: 10.1093/sysbio/syu036.

Jukes, Thomas H. and Charles R. Cantor (1969). "Evolution of protein molecules". In: *Mammalian protein metabolism* 3, pp. 21–132.

Kedzierska, Anna M. and Marta Casanellas (2012). "GenNon-h: Generating multiple sequence alignments on nonhomogeneous phylogenetic trees". In: *BMC Bioinformatics* 13(1), p. 216. ISSN: 1471-2105. DOI: 10.1186/1471-2105-13-216.

Kimura, Motoo (1980). "A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences". In: *Journal of Molecular Evolution* 16(2), pp. 111–120. ISSN: 0022-2844. DOI: 10.1007/BF01731581.

Kimura, Motoo (1981). "Estimation of evolutionary distances between homologous nucleotide sequences." In: *Proceedings of the National Academy of Sciences* 78(1), pp. 454–458. ISSN: 0027-8424. DOI: 10.1073/pnas.78.1.454.

Klaere, Steffen and Volkmar Liebscher (2012). "An algebraic analysis of the two state Markov model on tripod trees". In: *Mathematical Biosciences* 237(1), pp. 38–48. ISSN: 0025-5564. DOI: https://doi.org/10.1016/j.mbs.2012.03.001.

Kolaczkowski, Bryan and Joseph W. Thornton (2004). "Performance of maximum parsimony and likelihood phylogenetics when evolution is heterogeneous". In: *Nature* 431(7011), pp. 980–984. ISSN: 0028-0836. DOI: 10.1038/nature02917.

Kosta, Dimitra and Kaie Kubjas (2019). "Maximum Likelihood Estimation of Symmetric Group-Based Models via Numerical Algebraic Geometry". In: *Bulletin of Mathematical Biology* 81(2), pp. 337–360. ISSN: 0092-8240. DOI: 10.1007/s11538-018-0523-2.

Kreinin, Alexander and Marina Sidelnikova (2001). "Regularization algorithms for transition matrices". In: *Algo Research Quarterly* 4(1/2), pp. 23–40.

Kubatko, Laura S. and Julia Chifman (2019). "An invariants-based method for efficient identification of hybrid species from large-scale genomic data". In: *BMC Evolutionary Biology* 19(1), p. 112. ISSN: 1471-2148. DOI: 10.1186/s12862-019-1439-7.

Kück, Patrick et al. (2012). "Long Branch Effects Distort Maximum Likelihood Phylogenies in Simulations Despite Selection of the Correct Model". In: *PLoS ONE* 7(5). Ed. by John W. Stiller, e36593. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0036593.

Lake, James A. (1987). "A rate-independent technique for analysis of nucleic acid sequences: evolutionary parsimony." In: *Molecular Biology and Evolution* 4, pp. 167–191. ISSN: 1537-1719. DOI: 10.1093/oxfordjournals.molbev.a040433.

Lake, James A. (1994). "Reconstructing evolutionary trees from DNA and protein sequences: paralinear distances." In: *Proceedings of the National Academy of Sciences* 91(4), pp. 1455–1459. ISSN: 0027-8424. DOI: 10.1073/pnas.91.4.1455.

Lamarck, Jean Baptiste (1809). *Philosophie Zoologique*. Cambridge University Press.

Li, Tingting et al. (2020). "Phylogenetic supertree reveals detailed evolution of SARS-CoV-2". In: *Scientific Reports* 10(1), p. 22366. ISSN: 2045-2322. DOI: 10.1038/s41598-020-79484-8.

Margush, T and F R McMorris (1981). "Consensus n-trees". In: *Bulletin of Mathematical Biology* 43(2), pp. 239–244. ISSN: 0092-8240. DOI: 10.1007/BF02459446.

Matsen, Frederick A. (2009). "Fourier Transform Inequalities for Phylogenetic Trees". In: *IEEE/ACM Transactions on Computational Biology and Bioinformatics* 6(1), pp. 89–95. ISSN: 1545-5963. DOI: 10.1109/TCBB.2008.68.

Meyer, Carl D. (2000). *Matrix analysis and applied linear algebra*. Society for Industrial and Applied Mathematics: Philadelphia, PA, USA. ISBN: 0-89871-454-0.

Michelot, Cristian (1986). "A finite algorithm for finding the projection of a point onto the canonical simplex of $\propto^n$". In: *Journal of Optimization Theory and Applications* 50(1), pp. 195–200. ISSN: 0022-3239. DOI: 10.1007/BF00938486.

Mihaescu, Radu, Dan Levy, and Lior Pachter (2009). "Why Neighbor-Joining Works." In: *Algorithmica* 54(1), pp. 1–24.

Murphy, William J. et al. (2001). "Molecular phylogenetics and the origins of placental mammals". In: *Nature* 409(6820), pp. 614–618. ISSN: 0028-0836. DOI: 10.1038/35054550.

Pachter, Lior and Bernd Sturmfels (2005). *Algebraic Statistics for computational biology*. Ed. by L Pachter and B Sturmfels. Cambride University Press.

Paradis, Emmanuel, Julien Claude, and Korbinian Strimmer (2004). *APE: Analyses of Phylogenetics and Evolution in R language*. 2, pp. 289–290. DOI: 10.1093/bioinformatics/btg412.

Perron, Oskar (1907). “Zur Theorie der Matrices”. In: *Mathematische Annalen* 64(2), pp. 248–263. ISSN: 0025-5831. DOI: 10.1007/BF01449896.

Phillips, Matthew J., Frédéric Delsuc, and David Penny (2004). “Genome-Scale Phylogeny and the Detection of Systematic Biases”. In: *Molecular Biology and Evolution* 21(7), pp. 1455–1458. ISSN: 1537-1719. DOI: 10.1093/molbev/msh137.

Rambaut, Andrew and Nicholas C. Grass (1997). “Seq-Gen: an application for the Monte Carlo simulation of DNA sequence evolution along phylogenetic trees”. In: *Bioinformatics* 13(3), pp. 235–238. ISSN: 1367-4803. DOI: 10.1093/bioinformatics/13.3.235.

Ranwez, Vincent and Olivier Gascuel (2001). “Quartet-Based Phylogenetic Inference: Improvements and Limits”. In: *Molecular Biology and Evolution* 18(6), pp. 1103–1116. ISSN: 1537-1719. DOI: 10.1093/oxfordjournals.molbev.a003881.

Renard, Jean-Bruno (2018). “Jean-Loïc Le Quellec, Bernard Sergent (éds.), Dictionnaire critique de mythologie”. In: *Archives de sciences sociales des religions*(184), pp. 324–326. ISSN: 0335-5985. DOI: 10.4000/assr.45020.

Robinson, D.F. and L.R. Foulds (1981). “Comparison of phylogenetic trees”. In: *Mathematical Biosciences* 53(1-2), pp. 131–147. ISSN: 00255564. DOI: 10.1016/0025-5564(81)90043-2.

Roca-Lacostena, Jordi (2021). “The embedding problem for Markov matrices”. PhD thesis. UPC, Facultat de Matemàtiques i Estadística, Departament de Matemàtica Aplicada I.

Rokas, Antonis et al. (2003). “Genome-scale approaches to resolving incongruence in molecular phylogenies”. In: *Nature* 425(6960), pp. 798–804. ISSN: 0028-0836. DOI: 10.1038/nature02053.

Rosenberg, Michael S. and Sudhir Kumar (2001). “Incomplete taxon sampling is not a problem for phylogenetic inference”. In: *Proceedings of the National Academy of Sciences* 98(19), pp. 10751–10756. ISSN: 0027-8424. DOI: 10.1073/pnas.191248498.

Sabaté-Vidales, Marc (2014). “Reconstruction of phylogenetic trees using quartet methods”. PhD thesis. UPC, Facultat de Matemàtiques i Estadística, Departament de Matemàtica Aplicada I. URL: http://hdl.handle.net/2099.1/20844.

Saitou, Naruya and Masatoshi Nei (1987). “The neighbor-joining method: a new method for reconstructing phylogenetic trees.” In: *Molecular Biology and Evolution* 4(4), pp. 406–425. ISSN: 1537-1719. DOI: 10.1093/oxfordjournals.molbev.a040454.

Schottelius, Justus Georg (1995). *Ausführliche Arbeit von der teutschen HaubtSprache*. Ed. by Wolfgang Hecht. DE GRUYTER. ISBN: 978-3-484-16008-8. DOI: 10.1515/9783110940466.

Semple, Charles and Mike A. Steel (2003). *Phylogenetics*. Vol. 24. Oxford Lecture Series in Mathematics and its Applications. Oxford University Press: Oxford, pp. xiv+239. ISBN: 0-19-850942-1.

Snir, Sagi and Satish Rao (2010). "Quartets MaxCut: A Divide and Conquer Quartets Algorithm". In: *IEEE/ACM Transactions on Computational Biology and Bioinformatics* 7(4), pp. 704–718. ISSN: 1545-5963. DOI: 10.1109/TCBB.2008.133.

St. John, Katherine et al. (2003). "Performance study of phylogenetic methods: (unweighted) quartet methods and neighbor-joining". In: *Journal of Algorithms* 48(1), pp. 173–193. ISSN: 01966774. DOI: 10.1016/S0196-6774(03)00049-X.

Steel, Mike A. (2016). *Phylogeny: Discrete and Random Processes in Evolution*. SIAM-Society for Industrial and Applied Mathematics: Philadelphia, PA, USA. ISBN: 161197447X.

Steel, Mike A., Daniel Huson, and Peter J. Lockhart (2000). "Invariable Sites Models and Their Use in Phylogeny Reconstruction". In: *Systematic Biology* 49(2), pp. 225–232. ISSN: 1076-836X. DOI: 10.1093/sysbio/49.2.225.

Strimmer, Korbinian and Arndt von Haeseler (1996). "Quartet Puzzling: A Quartet Maximum-Likelihood Method for Reconstructing Tree Topologies". In: *Molecular Biology and Evolution* 13(7), pp. 964–969. ISSN: 0737-4038. DOI: 10.1093/oxfordjournals.molbev.a025664.

Strimmer, Korbinian and Arndt von Haeseler (1997). "Likelihood-mapping: A simple method to visualize phylogenetic content of a sequence alignment". In: *Proceedings of the National Academy of Sciences* 94(13), pp. 6815–6819. ISSN: 0027-8424. DOI: 10.1073/pnas.94.13.6815.

Sturmfels, Bernd and Seth Sullivant (2005). "Toric Ideals of Phylogenetic Invariants". In: *Journal of Computational Biology* 12(2), pp. 204–228. ISSN: 1066-5277. DOI: 10.1089/cmb.2005.12.204.

Sukumaran, Jeet and Mark T. Holder (2010). "DendroPy: a Python library for phylogenetic computing". In: *Bioinformatics* 26(12), pp. 1569–1571. ISSN: 1367-4803. DOI: 10.1093/bioinformatics/btq228.

Sumner, Jeremy G. et al. (2008). "Markov invariants, plethysms, and phylogenetics". In: *Journal of Theoretical Biology* 253(3), pp. 601–615. ISSN: 00225193. DOI: 10.1016/j.jtbi.2008.04.001.

Sumner, Jeremy G. et al. (2017). "Developing a statistically powerful measure for quartet tree inference using phylogenetic identities and Markov invariants". In: *Journal of Mathematical Biology* 75(6-7), pp. 1619–1654. ISSN: 0303-6812. DOI: 10.1007/s00285-017-1129-2.

Székely, László A.A., Mike A. Steel, and Péter L. Erdös (1993). "Fourier Calculus on Evolutionary Trees". In: *Advances in Applied Mathematics* 14(2), pp. 200–216. ISSN: 01968858. DOI: 10.1006/aama.1993.1011.

Tavaré, Simon (1986). "Some probabilistic and statistical problems on the analysis of DNA sequences". In: *Lectures on Mathematics in the Life Sciences* 17, pp. 57–86.

The Sage Developers (2019). *SageMath, the Sage Mathematics Software System (Version 8.6)*. URL: https://www.sagemath.org.

Verschelde, Jan (1999). "PHCpack: a general-purpose solver for polynomial systems by homotopy continuation Share on". In: *ACM Transactions on Mathematical Software* 25(2), pp. 251–276. ISSN: 0098-3500. DOI: 10.1145/317275.317286.

Warnow, Tandy (2017). *Computational Phylogenetics*. Cambridge University Press. ISBN: 9781107184718. DOI: 10.1017/9781316882313.

Willson, Stephen J. (1999). "Building Phylogenetic Trees from Quartets by Using Local Inconsistency Measures". In: *Molecular Biology and Evolution* 16(5), pp. 685–693. ISSN: 0737-4038. DOI: 10.1093/oxfordjournals.molbev.a026151.

Yang, Ziheng (1994). "Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: Approximate methods". In: *Journal of Molecular Evolution* 39(3), pp. 306–314. ISSN: 0022-2844. DOI: 10.1007/BF00160154.

Yang, Ziheng and Dave Roberts (1995). "On the use of nucleic acid sequences to infer early branchings in the tree of life." In: *Molecular Biology and Evolution*. ISSN: 1537-1719. DOI: 10.1093/oxfordjournals.molbev.a040220.

Zuckerkandl, Emile and Linus Pauling (1965). "Molecules as documents of evolutionary history". In: *Journal of Theoretical Biology* 8(2), pp. 357–366. ISSN: 00225193. DOI: 10.1016/0022-5193(65)90083-4.

Zwiernik, Piotr and Jim Q Smith (2011). "Implicit inequality constraints in a binary tree model". In: *Electronic Journal of Statistics* 5, pp. 1276–1312. ISSN: 1935-7524. DOI: 10.1214/11-EJS640.

A

CODES TO COMPUTE THE DISTANCE TO STOCHASTIC PHYLOGENETIC REGIONS

We present codes used for computations in Chapter 2. The following code in Magma (Bosma, Cannon, and Playoust, 1997) computes the EDdegree for the JC69 variety.

Code A.1: EDdegree of the Jukes Cantor variety for quartets. Code written in Magma.

```
R<x0, x1, x2, x3, x4> := PolynomialRing(Rationals(), 5, "grevlex");
p := 100; // It can be modified by the user;

u0 := Random(-p,p)/Random(-p,p); u1 := Random(-p,p)/Random(-p,p);
u2 := Random(-p,p)/Random(-p,p); u3 := Random(-p,p)/Random(-p,p);
u4 := Random(-p,p)/Random(-p,p);
f := 12(u0u1u2u3u4 - x0x1x2x3x4)^2 + 9(u0u1u2u3 - x0x1x2x3)^2 +
      6(u0u1u2u4 - x0x1x2x4)^2 + 6(u0u1u3u4 - x0x1x3x4)^2 +
      6(u0u2u3u4 - x0x2x3x4)^2 + 6(u1u2u3u4 - x1x2x3x4)^2 +
      3(u0u2u4 - x0x2x4)^2 + 3(u1u2u4 - x1x2x4)^2 +
      3(u0u3u4 - x0x3x4)^2 + 3(u1u3u4 - x1x3x4)^2 +
      3(u0u1 - x0x1)^2 + 3(u2u3 - x2x3)^2;
g := x0x1x2x3x4;
I := ideal<R|Derivative(f,x0),Derivative(f,x1),Derivative(f,x2),
             Derivative(f,x3),Derivative(f,x4)>;
J := ideal<R|g>;
S := Saturation(I,J);
P := ProjectiveSpace(Rationals(), 5); A := AffinePatch(P, 1);
X := Scheme(A, Generators(S));
print Degree(ProjectiveClosure(X));
```

The following code written in `Macaulay2` (Grayson and Stillman, 2009) has been used to test Conjecture 2.4.8 and do the simulations in Section 2.5.

Code A.2: Main code to test Conjecture 2.4.8 and perform simulations in Section 2.5
`AlgorithmStochasticPhylogeneticVarieties.m2`

```
restart
loadPackage "PHCpack";

localPath = currentFileDirectory;
load(concatenate{localPath, "functions.m2"});
tol = 1e-8;

- - - - PROBLEM DEFINITION - - - -

R = RR[x_0..x_4];
T = "1234"; - - Specify the phylogenetic variety V_T.

- - Load Fourier coordinates for points in V_T
M = lines get concatenate {localPath, concatenate{
                        "fourierCoordinates", T, ".txt"}};
MA = sub(value M_0, R); MC = sub(value M_0, R);

- - Random data point in the linear space L_T, see Example 1.3.20.
p = flatten entries random (R^1, R^14);
- - Uncomment the next two lines to test Conjecture 2.4.8.
- - k = random(0.,1.); m = random(1.,1.7);
- - p = {k,k,k^2,k^2,m,k*m,k*m,k*m,k^2*m,k^2*m,k*m,k^2*m,k^2*m,k^2*m};

- - Compute the objective funtion
MAp = matrix{{1,p_0,p_0,p_0},{p_1,p_2,p_3,p_3},
              {p_1,p_3,p_2,p_3},{p_1,p_3,p_3,p_2}};
MCp = matrix{{p_4,p_5,p_6,p_6},{p_7,p_8,p_9,p_9},
             {p_10,p_11,p_12,p_13},{p_10,p_11,p_12,p_13}};
param = flatten entries (MA - MAp | MC - MCp | MC - MCp | MC- MCp);
f = 0; for i to ♯param - 1 dof = f + param_i^2;
degJC69 = 249;

- - - - NON-SINGULAR CRITICAL POINTS AT THE INTERIOR - - - -

nonSingSolsList = NonSingSolsInterior(f, degJC69, tol);

- - - - CRITICAL POINTS AT THE BOUNDARY - - - -

boundaryList = CriticalPointsBoundary(R, f);

- - - - GLOBAL MINIMUM - - - -

L = join(nonSingSolsList, boundaryList);
(globMinPV, distPV) = FindingMinimum(f, nonSingSolsList);
(globMinSPR, distSPR) = FindingMinimum(f, L);

print(concatenate("Closest point to p in the phylogenetic variety V_T:",
      toString coordinates globMinPV, "; d(p,V_T)=" toString distPV));
print(concatenate("Closest point to p in the stochastic phylo region V_T^+ :",
      toString coordinates globMinSPR "; d(p, V_T^+)=" toString distSPR));
```

Code A.3: Supplementary functions for Code A.2. `functions.m2`

```
- - Round a complex number up to n decimals
roundCC = (z, n) -> (
    return round(n, realPart z) + round(n, imaginaryPart z)*ii;);

- - Round a all coordinates of a point up to n decimals
roundPointCC = (p, n) -> (
    return point {for i to ♯ p list(roundCC((coordinates p)_i, n))};);

- - Check if an element e is contained in a list L
inList = (L, e) -> (
    for i to ♯L - 1 do(if areEqual(L_i, e) then  return true;);
    return false;);

- - Returns a list with partial derivatives of f defined over CC
diffList = (f) -> (
    variables := support f;
    S := CC[variables];
    DF := diff(matrix {variables}, f);
    DFList := for  i from 0 to (numColumns DF - 1) list sub(DF_i_0, S);
    return DFList);

- - Returns real solutions with coordinates in [-1/3, 1]
validSolutions = (sols) -> (
    realSols := realPoints sols;
    validSols := for  i to ♯realSols-1 list(if not all (coordinates
    realSols_i, n -> n >= -1/3 and n <= 1) then  continue; realSols_i);
    return validSols );

- - Returns a point in CC^5.  Completes with the coordinates that are in
 the boundary
CompletePoint = (vars1, value1, vars2, value2, sol, gensR) -> (
    vars3 := toList (set gensR - set vars1 - set vars2);
    newPoint := new MutableList from gensR;
    for i to ♯vars1 - 1 do newPoint♯(index vars1_i) = value1;
    for i to ♯vars2 - 1 do newPoint♯(index vars2_i) = value2;
    for i to ♯vars3 - 1 do
       newPoint♯(index vars3_i) = roundCC((coordinates sol)_i, 12);
    return toList newPoint);

- - Returns a list with the border indices
BoundaryIndices = (variables) -> (
    BIaux = subsets variables;
    BI = for i from 0 to ♯BIaux - 1 list(
         {BIaux_i, subsets toList (set variables - set BIaux_i)});
    return BI;);

- - Given a function f and a list of points L = {x_1..x_n} it returns the
- - point x_i such that f(x_i) is minimum and the value f(x_i)
FindingMinimum = (f, L) -> (
    p = null; minVal = 1e10;
    for i to ♯L -1 do(
        val = (evaluate(polySystem{f}, point{L_i}))_(0,0);
        if sub(val,CC) < minVal then (
            minVal = val; p = point{L_i};
        );
    );
    return(p, minVal););
```

```
- - Given the polynomial f and the expected number of solutions degJC69 it
- - founds all roots of f and returns those that are on the feasible
 region.
NonSingSolsInterior = (f, degJC69, tol) -> (
    I = diffList(f);
    nonSingSolsList = new List;
    while ♯ nonSingSolsList < degJC69 do(
       sols = solveSystem(I);
       singSols = unique flatten for i to 4 list(zeroFilter(sols, i, tol));
       - - Filter non-sing solutions
       nonsingSols = select(sols, i-> not inList(singSols, i));
       nonsingSols = refineSolutions(I, nonsingSols, 64);
       nonsingSols = select(nonsingSols, i -> all(
                           flatten entries evaluate(polySystem I, i),
                           n -> abs(n) < tol));
       - - Add new non-sing solutions
       newSols = select(nonsingSols, i -> not inList(nonSingSolsList, i));
       nonSingSolsList = join(nonSingSolsList, newSols);
   );
   - - Filter solution at the interior of the feasible region
   nonSingSolsList = validSolutions(nonSingSolsList);
   return(nonSingSolsList);
);
- - Given a function f and a list of points L = {x_1..x_n} it returns the
- - point x_i such that f(x_i) is minimum and the value f(x_i)
CriticalPointsBoundary = (R, f) -> (
   L = new List;
   BI = BoundaryIndices(gens R);
   for i from 0 to ♯BI - 1 do(
       S1 = BI_i_0; - - Define set S1
       for j from 0 to ♯BI_i_1 - 1 do(
           S2 = BI_i_1_j; - Define set S2
           - - If both S1 and S2 are empty, then F = f
           if ♯S1 == 0 and ♯S2 == 0 then continue;
           F = sub(sub(f, for k to ♯S1-1 list(S1_k => 1)),
                          for k to ♯S2-1 list(S2_k => -1/3));
           if ♯S1 + ♯S2 < length gens R then (
             difF = diffList(F);
             sols = solveSystem(difF);
             sols = validSolutions(sols);
             if ♯sols == 0 then continue;
             feasibleBoundPoints = unique apply(sols, s ->
                                   CompletePoint(S1,1,S2,-1/3,s,gens R));
          )else feasibleBoundPoints = {CompletePoint(S1, 1, S2, -1/3, {},
                                        gens R)};
          L = join(L, feasibleBoundPoints);
      );
   );
   return(L);
);
```

B

Results of Q-methods. Additional tables

Here we present tables with the results of the three Q-methods Quartet-Puzzle, Willson and Weight optimization methods using the weighing systems given by `ASAQ`, `Erik+2`, `ML` and `NJ`, when applied to different data introduced in Section 4.1.2. We also include results for the global `NJ` with the same data sets.

The values on the 4th to 7th columns correspond to the average (for each of the 100 considered alignments) of the *Robinson-Foulds distance* from the original tree to the consensus tree obtained after 100 iterations of the Q-method. In case the quartet method failed to provide weights for some of the $\binom{12}{4}$ quartet subtrees, the corresponding alignment has been neglected.

The numbers in parentheses that appear in these tables represents the number of consensus trees that we have been able to reconstruct, if missing we have reconstructed the 100 trees. The value of the 8th column is the average value of its correspondent row. It indicates the average of the RF distance independently of the branch length parameter b. Finally, for each Q-method and each value of b, the minimum average RF distance is highlighted in bold type.

Q-METHODS ON GENERAL MARKOV DATA

Length (bp)	Q-method	weights	$b = 0.005$	$b = 0.015$	$b = 0.05$	$b = 0.1$	Average
600	QP	ASAQ	3.9	3.73	**5.38**	**8.44**	5.36
		Erik+2	**2.44**	**3.48**	6.16	9.63	5.43
		ML(homGMc)	9.55	9.58	9.67	10.67 (3)	9.87
		NJ	9.81	9.58	9.71	9.66	9.69
	WIL	ASAQ	**1.59**	**0.46**	**3.54**	**8.4**	3.5
		Erik+2	1.61	1.64	4.87	8.91	4.26
		ML(homGMc)	9.83	9.79	9.84	9 (3)	9.61
		NJ	8.34	7.96	8.83	9.45	8.64
	WO	ASAQ	**1.05**	**0.26**	**1.27**	**6.79**	2.34
		Erik+2	1.57	1.46	4.9	10.11	4.51
		ML(homGMc)	12.71	12.25	12.75	11 (3)	12.18
		NJ	10.31	9.77	10.1	10.92	10.28
	Global NJ		1.02	0.2	0.8	4.38	1.6
5000	QP	ASAQ	0.5	1.07	1.84	4.76	2.04
		Erik+2	**0.05**	**0.07**	**1.69**	**4.57**	1.6
		ML(homGMc)	9	9.08 (93)	9.65 (93)	- (0)	-
		NJ	9.7	9.65	9.91	9.91	9.79
	WIL	ASAQ	**0**	**0**	**0**	1.54	0.38
		Erik+2	0.02	0.07	0.33	**1.52**	0.48
		ML(homGMc)	9.97	9.28 (93)	9.27 (93)	- (0)	-
		NJ	7.66	7.41	8.25	9.61	8.23
	WO	ASAQ	**0**	**0**	**0**	**0.08**	0.02
		Erik+2	0.02	0.06	0.2	1.66	0.48
		ML(homGMc)	11.03	11.2 (93)	11.8 (93)	- (0)	-
		NJ	9.46	9.22	9.2	9.73	9.4
	Global NJ		0	0	0	0.14	0.04
10000	QP	ASAQ	0,04	0,28	1,37	3,2	1,22
		Erik+2	0	0	0,44	2,92	0,84
		ML(homGMc)	9	9,11	9,6	9	9,18
		NJ	9,69	9,97	9,92	9,95	9,88
	WIL	ASAQ	0	0	0	0,19	0,05
		Erik+2	0	0,02	0,09	0,44	0,14
		ML(homGMc)	9,33	9	8,98	9	9,08
		NJ	7,46	7,44	8,03	9,48	8,1
	WO	ASAQ	0	0	0	0	0
		Erik+2	0	0	0,08	0,51	0,15
		ML(homGMc)	10,97	11,01	11,69	13	11,67
		NJ	9,41	9,18	9,33	9,62	9,38
	Global NJ		0	0	0	0.04	0.01

TABLE B.1: Average Robinson-Foulds distance for GM data simulated on the tree *CC*. The number in parentheses indicates the number of consensus trees that have been reconstructed. If missing, all 100 trees have been inferred.

Length (bp)	Q-method	weights	$b = 0.005$	$b = 0.015$	$b = 0.05$	$b = 0.1$	Average
600	QP	ASAQ	6.89	7.87	8.65	10.1	8.38
		Erik+2	**5.9**	**6.92**	**8.23**	**9.96**	7.75
		ML(homGMc)	9.53 (78)	9.54 (98)	9.67 (95)	10 (1)	9.69
		NJ	10.09	9.81	10	10.25	10.04
	WIL	ASAQ	**4.68**	**4.43**	**6.09**	**9.51**	6.18
		Erik+2	4.82	5	6.84	9.79	6.61
		ML(homGMc)	9.87 (78)	9.9 (98)	9.61 (95)	9 (1)	9.59
		NJ	9.06	9.21	9.82	9.88	9.49
	WO	ASAQ	**4.4**	**4.22**	**4.88**	**8.47**	5.49
		Erik+2	4.83	4.93	7.72	11.34	7.21
		ML(homGMc)	12.28 (78)	12.04 (98)	12.32 (95)	10 (1)	11.66
		NJ	11.31	11.01	11.09	11.45	11.21
	Global NJ		4.42	4.2	4.34	7.70	5.17
5000	QP	ASAQ	4.56	5.11	5.81	8.65	6.03
		Erik+2	**4.03**	**4.09**	**5.5**	**7.71**	5.33
		ML(homGMc)	9	9.08 (96)	9.58 (96)	10 (1)	9.41
		NJ	9.65	9.88	9.9	10.44	9.97
	WIL	ASAQ	**4**	**4**	**4**	5.01	4.25
		Erik+2	4.02	4.06	4.24	**4.97**	4.32
		ML(homGMc)	10	9.34 (96)	9.34 (96)	9 (1)	9.42
		NJ	8.92	9.12	9.43	9.85	9.33
	WO	ASAQ	**4**	**4**	**4**	**4.06**	4.01
		Erik+2	**4**	4.04	4.16	5.38	4.39
		ML(homGMc)	11.04	11.18 (96)	11.76 (96)	10 (1)	10.99
		NJ	10.64	10.62	10.62	10.75	10.66
	Global NJ		4	4	4	4.12	4.03
10000	QP	ASAQ	4,05	4,3	5,39	7,26	5,25
		Erik+2	4	4	4,44	6,76	4,8
		ML(homGMc)	9	9,09	9,6	10	9,42
		NJ	9,59	9,79	10,3	10,2	9,97
	WIL	ASAQ	4	4	4	4,13	4,03
		Erik+2	4	4,02	4,08	4,35	4,11
		ML(homGMc)	9,33	8,99	8,99	9	9,08
		NJ	9,01	9,11	9,76	9,94	9,45
	WO	ASAQ	4	4	4	4	4
		Erik+2	4	4	4,04	4,39	4,11
		ML(homGMc)	10,96	10,96	11,55	9	10,62
		NJ	10,76	10,54	10,83	10,81	10,73
	Global NJ		4	4	4	4.04	4.01

TABLE B.2: Average Robinson-Foulds distance for GM data simulated on the tree *CD*. The number in parentheses indicates the number of consensus trees that have been reconstructed. If missing, all 100 trees have been inferred.

Length (bp)	Q-method	weights	$b = 0.005$	$b = 0.015$	$b = 0.05$	$b = 0.1$	Average
600	QP	ASAQ	10.62	11.4	12.32	11.29	11.41
		Erik+2	9.29	10.16	10.07	10.36	9.97
		ML(homGMc)	9.47 (74)	9.6 (93)	9.73 (89)	9.62 (8)	9.61
		NJ	**8.7**	**8.85**	**8.91**	**9.21**	8.92
	WIL	ASAQ	**8.19**	**8.38**	9.26	10.26	9.02
		Erik+2	8.43	8.49	**9.1**	10.24	9.07
		ML(homGMc)	9.88 (74)	9.72 (93)	9.6 (89)	9.38 (8)	9.65
		NJ	8.91	8.9	9.59	**9.31**	9.18
	WO	ASAQ	**7.99**	**8.06**	**8.78**	**9.59**	8.61
		Erik+2	8.39	8.47	9.67	12	9.63
		ML(homGMc)	12.26 (74)	11.92 (93)	11.89 (89)	11 (8)	11.77
		NJ	11.11	11.06	11.18	11.29	11.16
	Global NJ		8.38	8.10	8.20	10.41	8.74
5000	QP	ASAQ	8.52	8.99	9.68	12.32	9.88
		Erik+2	**8.03**	**8.09**	9.36	11.11	9.15
		ML(homGMc)	9	9.11 (85)	9.59 (87)	- (0)	-
		NJ	9.11	8.95	**8.97**	**9**	9.01
	WIL	ASAQ	**8**	**8**	**8**	8.54	8.13
		Erik+2	8.02	**8**	8.04	**8.33**	8.1
		ML(homGMc)	10.04	9.31 (85)	9.36 (87)	- (0)	-
		NJ	8.82	8.81	9.54	9.31	9.12
	WO	ASAQ	**8**	8	**8**	**8.04**	8.01
		Erik+2	**8**	**7.99**	8.02	8.68	8.17
		ML(homGMc)	11.04	11.18 (85)	11.79 (87)	- (0)	-
		NJ	10.65	10.52	10.5	10.79	10.62
	Global NJ		8	8	8	8.04	8.01
10000	QP	ASAQ	8,08	8,29	9,35	11,33	9,26
		Erik+2	8	8	8,48	10,24	8,68
		ML(homGMc)	9	9,08	9,64	0	0
		NJ	9,09	8,91	9,12	9,26	9,09
	WIL	ASAQ	8	8	8	8,04	8,01
		Erik+2	8	8	7,94	8	7,99
		ML(homGMc)	9,34	8,97	8,99	0	0
		NJ	8,82	8,8	9,53	9,84	9,25
	WO	ASAQ	8	8	8	8	8
		Erik+2	8	8	7,96	8,03	8
		ML(homGMc)	10,98	10,91	11,52	0	0
		NJ	10,71	10,62	10,74	10,89	10,74
	Global NJ		8	8	8	8.04	8.01

TABLE B.3: Average Robinson-Foulds distance for GM data simulated on the tree *DD*. The number in parentheses indicates the number of consensus trees that have been reconstructed. If missing, all 100 trees have been inferred.

Q-METHODS ON GENERAL TIME REVERSIBLE DATA

Length (bp)	Q-method	weights	$b = 0.005$	$b = 0.015$	$b = 0.05$	$b = 0.1$	Average
600	QP	ASAQ	4.06	5	6.91	10.44	6.6
		Erik+2	**3.7**	**4.76**	**4.89**	**8.35**	5.43
		ML(homGMc)	10.1 (77)	9.88 (99)	9.65 (97)	- (0)	-
		NJ	9.18	9.05	9.19	9.54	9.24
	WIL	ASAQ	**1.9**	**1.11**	8.64	11.58	5.81
		Erik+2	2.41	2.43	**3.33**	**7.75**	3.98
		ML(homGMc)	9.71 (77)	9.57 (99)	9.43 (97)	- (0)	-
		NJ	8.42	8.23	8.46	9.25	8.59
	WO	ASAQ	**1.48**	**0.37**	3.28	12.54	4.42
		Erik+2	2.65	2.04	**2.3**	**7.15**	3.54
		ML(homGMc)	12.71 (77)	12.45 (99)	11.98 (97)	- (0)	-
		NJ	11.19	10.93	11.18	11.12	11.1
	Global NJ		1.32	0.20	0.92	5.40	1.96
5000	QP	ASAQ	2.78	3.6	**3.77**	**5.02**	3.79
		Erik+2	**0.83**	**2.69**	3.91	5.22	3.16
		ML(homGMc)	9.19 (64)	9.28 (86)	9.28 (97)	9.82 (33)	9.39
		NJ	8.93	8.89	9	9.26	9.02
	WIL	ASAQ	**0**	**0**	**0**	5.79	1.45
		Erik+2	0.1 (99)	**0**	0.08	**0.72**	0.22
		ML(homGMc)	9.6 (62)	9.29 (86)	9.38 (97)	9.3 (33)	9.39
		NJ	7.87 (99)	7.93	7.97	8.69	8.12
	WO	ASAQ	**0**	**0**	**0**	0.97	0.24
		Erik+2	**0**	**0**	0.02	**0.45**	0.12
		ML(homGMc)	11.88 (64)	11.81 (86)	12.04 (97)	11.88 (33)	11.9
		NJ	10.82	10.78	11.3	11.73	11.16
	Global NJ		0	0	0	0.38	0.1

TABLE B.4: Average Robinson-Foulds distance for GTR data simulated on the tree *DD*. The number in parentheses indicates the number of consensus trees that have been reconstructed. If missing, all 100 trees have been inferred.

Q-METHODS ON MIXTURE DATA

Length (bp)	Q-method	weights	$b = 0.005$	$b = 0.015$	$b = 0.05$	$b = 0.1$	Average
600	QP	ASAQ	3.91	4.01	5.03	7.95 (98)	5.22
		ASAQ (m=2)	**1.93**	3.7	4.59	7.48	4.42
		Erik+2	3.54	**3.23**	**4.19**	7.36	4.58
		ML(homGMc)	9.66	10.24	9.65	9 (1)	9.64
		NJ	5.71	5.79	5.37	**6.46**	5.83
	WIL	ASAQ	**1.96**	0.96	3.14	7.46 (98)	3.38
		ASAQ (m=2)	3.04	**0.88**	**2.26**	**6.88**	3.27
		Erik+2	4	1.57	2.71	6.98	3.82
		ML(homGMc)	10.07	9.57	9.49	9 (1)	9.53
		NJ	6.18	6.31	6.59	8.09	6.79
	WO	ASAQ	1.74	**0.46**	1.73	6.68 (98)	2.65
		ASAQ (m=2)	**1.63**	0.48	**1.24**	**5.8**	2.29
		Erik+2	3.49	1.36	2.1	6.94	3.47
		ML(homGMc)	12.84	12.1	11.84	12 (1)	12.2
		NJ	9.02	8.36	8.24	9.61	8.81
	Global NJ		1.84	0.16	0.96	4.98	1.99
5000	QP	ASAQ	1.55	2.08	2.93	5.11	2.92
		ASAQ (m=2)	0.88	2.24	2.59	5.02	2.68
		Erik+2	**0.14**	**0.63**	**2.33**	**4.33**	1.86
		ML(homGMc)	9.11	9.29	9.87	9 (1)	9.32
		NJ	4.95	4.76	4.74	5.75	5.05
	WIL	ASAQ	**0**	**0**	0.3	2.3	0.65
		ASAQ (m=2)	**0**	**0**	0.35	2.3	0.66
		Erik+2	0.04	0.08	**0.08**	**1.3**	0.38
		ML(homGMc)	9.95	9.53	9.32	10 (1)	9.7
		NJ	5.91	5.24	5.71	7.78	6.16
	WO	ASAQ	**0**	**0**	**0**	0.95	0.24
		ASAQ (m=2)	**0**	**0**	**0**	0.96	0.24
		Erik+2	0.06	0.02	0.04	**0.73**	0.21
		ML(homGMc)	11.32	11.21	11.29	10 (1)	10.96
		NJ	9.16	8.24	7.51	8.85	8.44
	Global NJ		0	0	0.06	1.68	0.44
10000	QP	ASAQ	0.88	1.43	2.54	4.47	2.33
		ASAQ (m=2)	0.96	1.76	2.47	4.7	2.47
		Erik+2	**0.04**	**0.08**	**1.52**	**3.69**	1.33
		ML(homGMc)	9.28	9.15	9.8	10.25 (4)	9.62
		NJ	5	4.7	4.73	5.64	5.02
	WIL	ASAQ	**0**	**0**	0.16	1.92	0.52
		ASAQ (m=2)	**0**	**0**	0.19	2.49	0.67
		Erik+2	0.02	**0**	**0.06**	**0.83**	0.23
		ML(homGMc)	9.24	8.98	9.03	9 (4)	9.06
		NJ	5.98	5.59	5.76	7.75	6.27
	WO	ASAQ	**0**	**0**	**0.02**	1.31	0.33
		ASAQ (m=2)	**0**	**0**	**0.02**	0.98	0.25
		Erik+2	0.02	**0**	0.04	**0.58**	0.16
		ML(homGMc)	11	11.07	11.27	10.75 (4)	11.02
		NJ	9.07	8.85	7.15	8.7	8.44
	Global NJ		0	0	0.04	1.64	0.42

TABLE B.5: Average Robinson-Foulds distance for 2-category data with $p = 0.25$. The number in parentheses indicates the number of consensus trees that have been reconstructed. If missing, all 100 trees have been inferred.

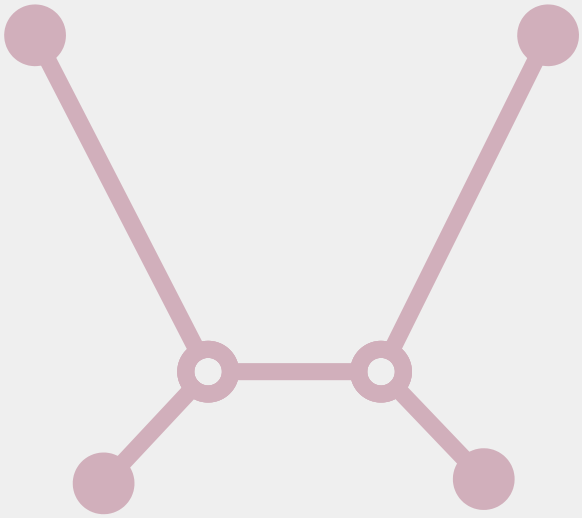
Length (bp)	Q-method	weights	$b = 0.005$	$b = 0.015$	$b = 0.05$	$b = 0.1$	Average
600	QP	ASAQ	4.56	4.46	5.15	7.57 (99)	5.43
		ASAQ (m=2)	**3**	4.04	4.85	**7.06**	4.74
		Erik+2	4.89	**3.88**	**4.54**	7.21	5.13
		ML(homGMc)	9.67	9.89	9.7	10 (3)	9.82
		NJ	7.05	6.72	7.02	7.42	7.05
	WIL	ASAQ	**3.29**	1.43	3.46	**6.27** (99)	3.61
		ASAQ (m=2)	3.99	**1.05**	**2.92**	6.3	3.56
		Erik+2	5.3	2.55	3.17	6.58	4.4
		ML(homGMc)	9.85	9.67	9.74	9 (3)	9.56
		NJ	7.23	6.89	7.28	8.02	7.36
	WO	ASAQ	**2.95**	**0.86**	1.83	5.58 (99)	2.8
		ASAQ (m=2)	3	**0.86**	**1.3**	**4.89**	2.51
		Erik+2	5.08	2.56	2.57	6.29	4.12
		ML(homGMc)	12.18	12.13	12.57	11.67 (3)	12.14
		NJ	9.23	8.43	8.62	9.33	8.9
	Global NJ		2.98	0.82	1.18	6.18	2.79
5000	QP	ASAQ	1.43	2.35	3.28	5.75	3.2
		ASAQ (m=2)	0.97	2.37	2.77	5.4	2.88
		Erik+2	**0.13**	**0.92**	**2.52**	**4.76**	2.08
		ML(homGMc)	9.25	9.38	10.07	10 (1)	9.68
		NJ	5.81	5.83	7	7.79	6.61
	WIL	ASAQ	**0**	**0**	0.8	3.27	1.02
		ASAQ (m=2)	0.04	**0**	0.67	3.5	1.05
		Erik+2	0.1	0.02	**0.25**	**1.7**	0.52
		ML(homGMc)	9.82	9.33	9.4	9 (1)	9.39
		NJ	6.02	6.49	6.77	7.9	6.8
	WO	ASAQ	**0**	**0**	0.16	1.81	0.49
		ASAQ (m=2)	**0**	**0**	**0.1**	1.24	0.34
		Erik+2	0.1	**0**	**0.1**	**1.08**	0.32
		ML(homGMc)	11.18	11.28	11.64	11 (1)	11.28
		NJ	8.76	8.44	8.34	8.77	8.58
	Global NJ		0	0	0.18	1.84	0.51
10000	QP	ASAQ	0.9	1.53	3.24	5.85	2.88
		ASAQ (m=2)	0.76	1.77	2.62	4.99	2.54
		Erik+2	**0.02**	**0.16**	**1.6**	**4.04**	1.46
		ML(homGMc)	9.13	9.17	9.63	9 (1)	9.23
		NJ	5.35	5.8	6.69	7.61	6.36
	WIL	ASAQ	**0**	**0**	0.7	2.97	0.92
		ASAQ (m=2)	**0**	**0**	0.57	3.16	0.93
		Erik+2	**0**	**0**	**0.02**	**0.85**	0.22
		ML(homGMc)	9.19	9.11	9	9 (1)	9.07
		NJ	6.28	5.92	6.4	7.98	6.65
	WO	ASAQ	**0**	**0**	0.2	2.28	0.62
		ASAQ (m=2)	**0**	**0**	0.07	1.45	0.38
		Erik+2	0.02	**0**	**0.02**	**0.42**	0.12
		ML(homGMc)	11.1	11.15	11.37	11 (1)	11.15
		NJ	8.33	8.39	8.27	8.93	8.48
	Global NJ		0	0	0.14	2.14	0.57

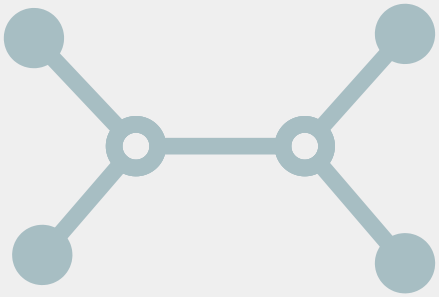
TABLE B.6: Average Robinson-Foulds distance for 2-category data with $p = 0.50$. The number in parentheses indicates the number of consensus trees that have been reconstructed. If missing, all 100 trees have been inferred.

Length (bp)	Q-method	weights	$b = 0.005$	$b = 0.015$	$b = 0.05$	$b = 0.1$	Average
600	QP	ASAQ	4.15	4.35	4.97	8.13	5.4
		ASAQ (m=2)	**1.88**	4.02	5.09	7.77	4.69
		Erik+2	4.35	**3.08**	**4.24**	**7.36**	4.76
		ML(homGMc)	9.88	9.67	9.64	- (0)	-
		NJ	9.14	9.04	8.9	9.09	9.04
	WIL	ASAQ	**2.07**	0.99	3.63	7.9	3.65
		ASAQ (m=2)	2.87	**0.77**	2.56	7.15	3.34
		Erik+2	5.12	1.74	**2.41**	**6.75**	4
		ML(homGMc)	9.81	9.89	9.36	- (0)	-
		NJ	7.59	7.31	8.12	9.17	8.05
	WO	ASAQ	1.82	0.87	1.57	6.27	2.63
		ASAQ (m=2)	**1.76**	**0.44**	**0.95**	**5.55**	2.17
		Erik+2	4.62	1.7	2.06	6.87	3.81
		ML(homGMc)	12.81	12.2	12.15	- (0)	-
		NJ	9.61	9.77	10.1	10.2	9.92
	Global NJ		1.46	0.6	1.12	5.4	2.15
5000	QP	ASAQ	1.39	1.99	2.91	5.47	2.94
		ASAQ (m=2)	0.88	2.36	2.77	5.34	2.84
		Erik+2	**0.1**	**0.58**	**2.52**	**4.81**	2
		ML(homGMc)	9.06	9.2	10.14	- (0)	6.85
		NJ	9.54	8.83	9.31	8.97	9.16
	WIL	ASAQ	**0**	**0**	0.16	3.12	0.82
		ASAQ (m=2)	**0**	**0**	0.3	3.21	0.88
		Erik+2	0.02	0.01	**0.07**	**1.45**	0.39
		ML(homGMc)	9.94	9.53	9.46	- (0)	-
		NJ	7.16	7.09	6.97	7.84	7.26
	WO	ASAQ	**0**	**0**	**0**	1.44	0.36
		ASAQ (m=2)	**0**	**0**	0.04	**1.01**	0.26
		Erik+2	0.06	**0**	0.04	1.08	0.3
		ML(homGMc)	12.11	11.43	11.68	- (0)	-
		NJ	9.62	9.52	9.21	9.14	9.37
	Global NJ		0	0	0.12	1.46	0.4
10000	QP	ASAQ	0.73	1.49	2.73	5.63	2.64
		ASAQ (m=2)	0.81	1.83	2.46	5	2.52
		Erik+2	**0**	**0.04**	**1.98**	**4.05**	1.52
		ML(homGMc)	9.06	9.19	9.73	- (0)	-
		NJ	9.11	9.27	9.37	9	9.19
	WIL	ASAQ	**0**	**0**	0.32	2.87	0.8
		ASAQ (m=2)	**0**	**0**	0.22	2.93	0.79
		Erik+2	**0**	0.02	**0.01**	**0.33**	0.09
		ML(homGMc)	9.09	8.86	9.03	- (0)	-
		NJ	7.27	6.57	6.7	8.39	7.23
	WO	ASAQ	**0**	**0**	0.12	1.54	0.42
		ASAQ (m=2)	**0**	**0**	0.08	0.81	0.22
		Erik+2	**0**	0.02	**0**	**0.28**	0.08
		ML(homGMc)	11.25	11.28	11.41	- (0)	-
		NJ	9.4	8.73	9.21	9.73	9.27
	Global NJ		0	0	0.04	1.74	0.45

TABLE B.7: Average Robinson-Foulds distance for 2-category data with $p = 0.75$. The number in parentheses indicates the number of consensus trees that have been reconstructed. If missing, all 100 trees have been inferred.

LaTeXstyle: Based on the Masters/Doctoral Thesis template Version 2.5
by Vel, originally created by Steve Gunn and Sunil Patel.





UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONA**TECH**